



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 142280**

**To: Barba Koroma**  
**Location: REM/2C18**  
**Art Unit: 1638**  
**Wednesday, January 26, 2005**

**Case Serial Number: 10/808979**

**From: Beverly Shears**  
**Location: Remsen Bldg.**  
**RM 1A54**  
**Phone: 571-272-2528**

**beverly.shears@uspto.gov**

### **Search Notes**

**This Page Blank (uspto)**

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 24, 2005, 23:06:56 ; Search time 28265 Seconds  
(without alignments)  
11709.910 Million cell updates/sec

Title: US-10-808-979-18  
Perfect score: 6999  
Sequence: 1 gcgatggcaatggaaatgt.....gacgttttagacatgcaata 6999

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2732.8	39.0	6815	AB076662	AB076662 Pantoea a
2	2626.2	37.5	6918	1 ERWCRT	D90087 Erwinia ure
3	2624.6	37.5	6918	6 I12910	I12910 Sequence 13
4	2620	37.4	6586	1 AY166713	AY166713 Pantoea s
5	2468.8	35.3	6965	1 ERWCRTA	M90698 Pantoea agg
6	2135.2	30.5	12753	1 ERWCRTS	M87280 Pantoea agg
7	1900.6	27.2	4624	6 CQ93208	CQ93208 Sequence
8	1374.8	19.6	349107	1 BX571873	BX571873 Photorhab
9	1189	17.0	2415	1 ERWCRTIB	M38423 Pantoea agg
10	1179.8	16.9	3801	1 AF408848	AF408848 Xanthobac
11	1105.8	15.8	6635	1 AF218415	AF218415 Bradyrhiz
12	1041.2	14.9	6335	1 PMCRTWZYI	Y15112 Paracoccus
13	1037.2	14.8	5373	1 ATUCRTWA	D58420 Paracoccus
14	1035.4	14.8	5188	1 FAU62808	U62808 Flavobacter
15	1035.4	14.8	8625	6 AR169831	AR169831 Sequence
16	1035.4	14.8	8625	6 AR452005	AR452005 Sequence
17	1035.4	14.8	11233	6 AR169852	AR169852 Sequence
18	1035.4	14.8	11233	6 AR452026	AR452026 Sequence
19	827.4	11.8	1493	6 BD241820	BD241820 Method of

20	827.4	11.8	1493	6	AX014697	AX014697 Sequence
21	826.2	11.8	1479	6	CQ793426	CQ793426 Sequence
22	826.2	11.8	1479	6	CQ793595	CQ793595 Sequence
23	826.2	11.8	1479	6	E22381	E22381 beta-Carote
24	826.2	11.8	1479	6	I12907	I12907 Sequence 10
25	808.6	11.6	1479	6	AX394988	AX394988 Sequence
26	713.4	10.2	1518	6	I72653	I72653 Sequence 7
27	708.4	10.1	1522	6	I72654	I72654 Sequence 9
28	619.4	8.8	1482	6	A58565	A58565 Sequence 6
29	619.4	8.8	1482	6	A84702	A84702 Sequence 12
30	619.4	8.8	1482	6	AR103081	AR103081 Sequence
31	619.4	8.8	1482	6	AR139697	AR139697 Sequence
32	619.4	8.8	1482	6	AR391876	AR391876 Sequence
33	604.4	8.6	349723	1	BX842850	BX842850 Bdellovib
34	515.6	7.4	1198	6	I24838	I24838 Sequence 5
35	515.6	7.4	1198	6	I72652	I72652 Sequence 5
36	479	6.8	30040	1	AF182374	AF182374 Bradyrhiz
37	467.4	6.7	349737	1	BX572597	BX572597 Rhodopseu
38	461.4	6.6	1581	1	D83514	D83514 Erythrobact
39	461.4	6.6	1581	6	E11103	E11103 Gene coding
40	449.6	6.4	1083	6	I40224	I40224 Sequence 1
41	449.6	6.4	5397	12	AY605097	AY605097 Synthetic
42	449.2	6.4	1296	6	AX394984	AX394984 Sequence
43	438	6.3	1235	6	I72656	I72656 Sequence 13
44	430	6.1	1232	6	AR222522	AR222522 Sequence
45	430	6.1	1232	6	AR432935	AR432935 Sequence

## ALIGNMENTS

RESULT 1	AB076662	AB076662	6815 bp	DNA	linear	BCT 26-DEC-2001
LOCUS	Pantoea agglomerans pv. milletiae crtX, crtY, crtB,					
DEFINITION	crtZ genes, complete cds.					
ACCESSION	AB076662					
VERSION	AB076662.1	GI:18143444				
KEYWORDS						
SOURCE	Pantoea agglomerans pv. milletiae					
ORGANISM	Pantoea agglomerans pv. milletiae					
REFERENCE	1					
AUTHORS	Kamiunten,H. and Hirata,R.					
TITLE	Isolation and characterization of carotenoid biosynthesis genes					
JOURNAL	from Pantoea agglomerans pv. milletiae Wist 801					
REFERENCE	2 (bases 1 to 6815)					
AUTHORS	Kamiunten,H. and Hirata,R.					
TITLE	Direct Submission					
JOURNAL	Submitted (20-DEC-2001) Hiroshi Kamiunten, Miyazaki University,					
FACTORS	Faculty of Agriculture; Gakuenkibanadai Nishi 1-1, Miyazaki,					
source	Miyazaki 889-2155, Japan (E-mail: a01108@cc.miyazaki-u.ac.jp,					
Location/Qualifiers	Tel:81-985-58-7170, Fax:81-985-58-7170)					
1. .6815	/organism="Pantoea agglomerans pv. milletiae"					
/mol_type="genomic DNA"						
/db_xref="taxon:182454"						
/note="pathovar: milletiae"						
160. .1071	/gene="crtE"					
160. .1071	/gene="crtE"					
/codon_start=1						
/transl_table=11						
/protein_id="BAB79600.1"						
/db_xref="GI:18143445"						
/translation="MMTCAEONFTHSDAASLLNDIEQRDLQPLPVSERDLVGA						
MDGALAPGKRIRPLLLALLALRDLCGNATPAGLLDACAVEMVHAASLLIDDDPCMD						
QQLRGRTTHCQYGEHVAILAAVALLSKAFGVAAAEGLTATARADVAELSHAYGM						
QQLVQGFQDUSEGDKPRSDAILMTNHYKTTLFCASMQMASIAVAEASGEAREQLHR						

FSNLQAFQQLDDLDGMDATGKDAHQDDGKSTLVLNLLGPQAVETRLRLHLRCASEH  
 LLSACODGYATHFVQAWPEKKLAVS"

gene

1081..2376

CDS

/gene="ctx"

1081..2376

/gene="ctx"

/codon\_start=1

/transl\_table=1

/protein\_id="BAB79601.1"

/db\_xref="GI:18143446"

/translation="MSHFAVIAPPPFVSHVQALQHLISQALIARGHQITFTIQQDTVSALL

TSURIGFFPGLGSHAPGSLAHTLOLAHPLGLPSMLKLEINARSTDMLCRELPTVLS

TLAIDGVIYDQEPAGALAAEALNLPYVSACALPLNREADPFLAVMPDYARTDQAR

ERYTSEKIYDMLRRHRIYIARNHAMGLAPREKLHCFSPVAOINQILPELDPRQ

TLPDPHSPGLRTTTPAASQOPRYFPHGDRPRIIPASIGTIQGHRYGIFKTIIVRACQ

EIDAQLLAHCGRLSPFAEKLAQASHVQVDFADQAAALAAQADLVIHGGMNVTLDG

INHLPTLLIPIAFDOPGVAAVVMHGIQRASRFTTSHSMARQJOTLLADESYVOOM

KILRSALRQAGGTTAAADIVEQAMUTROPVLTRHYAAV"

2363..3523

gene

/gene="ctx"

2363..3523

/gene="ctx"

/codon\_start=1

/transl\_table=1

/protein\_id="BAB79602.1"

/db\_xref="GI:18143447"

/translation="MPRYDLILVAGLANGLIALRLROORPSLRILLIDAEREPCANH

TWSFHAEDTETQHRWIAPLVHVHWPGEVREPORSRNSNGVFCVTAERFAQVIRDR

FAPDILLNTRVSIASRAVTLDDGRVLESDAVIDGRVQPDGALRWGFSFVQGEWL

SPHGUTAPIIMDATVDQAGRYFVSILPFSFADTILLIEDHYIDNLEGDRARQNIH

DYAQQGRLDLREEGEALPITLIGDVAAFWONRDLPCSLGRAGLPHFTTGYSLPL

AVALLADRAQMTFTSETLHATIQFASQAWQOQRFRLNLRMLFLAGPADORWQVQ

RYGLPEGLIARFYAGKLTLPDLRLILSKPPVPLVLAALQALMTPHRQAMQ"

3520..4998

gene

/gene="ctx"

3520..4998

/gene="ctx"

/codon\_start=1

/transl\_table=1

/protein\_id="BAB79603.1"

/db\_xref="GI:18143448"

/translation="WNRITVIGAGFGLALAIRLOASGVPTLLEQRDKPGGRATVYQ

DQGFIDAGFVITDPSALBELFLTAGKKLSYDELMPVKPFPYRLCSWKGKVSFYDND

QPALBAQAFNPVREYFAYRAVFAEYGLKGLTVPFLSRDMLRAAPQLAKL

QAWRSYKVASYIDEHLRQAFSHSLVGNPATSIYTLIHAREWGVMPRG

GTGALVQGVKLFEDLQGEVNLASVARLEQENRITAVHLKDGGRVETRAVASNADY

VHTYRELLSOHPASOAGRSLOKEMSNLSFVIFGLAHNHNOLAHHTVCGPRYREL

IDEIIPKGLADEFSLYLHAPCVTPDPSLAPEGCGSYVLAEPVHLGLTADIDWAVEGR

LDRIFDIEQHYMPGLRSQVLTFRIFTPDFRDELNAYQGSFAFSVEPILTQSAWFRP

HNRDKNLNLVLVGATHTPGAGIPGIVGSAKATAGLMLDLA"

5034..5924

gene

/gene="ctx"

5034..5924

/gene="ctx"

/codon\_start=1

/transl\_table=1

/protein\_id="BAB79604.1"

/db\_xref="GI:18143449"

/translation="NEVGSKSPATSKLPGAKTRSLVLMYAWCEHDDVDVQVLF

SNDTPSLOSARQALQLEMKTRQAGSOMHEPAPAAQOEVMAMHDIILPAYAFDLHAG

FAMDVHETRYQDLTLRYCYTHVAGVGLMMAQIMGVRDNATLDRACDLGLAFQITNI

ARDIVEDADAGRCYLPAFWLAEGLTRENLDAPQNRQALSRRVRLVETAEPPYRSAS

AGLPGILRSAWAIATAOQVYRKIGMKVQAASQAWDQORQSTSTPEKLALLVAASQA

VTSRVARHAPRSADLWQRPV"

complement(5863..6390)

/gene="ctx"

complement(5863..6390)

/gene="ctx"

/codon\_start=1

/transl\_table=1

/protein\_id="BAB79605.1"

/db\_xref="GI:18143450"

/translation="MLMTWNLILVLTVIGMEITAAALAHRYIMHGWGWGHLSHHEPH

KGWFEVNDLYAVVFAALSILLIYLSTGVWPIOWIGAGMTLYGLLYFIVHDGLVHORW  
 PERYPVPRGYLRRLWYAHMHHAHVKGEGCVSFRFLYAPPLSKLQATLRRHGVKRGAA  
 ARDQSRVERDAPPGK"

ORIGIN

Query Match 39.0%; Score 2732.8; DB 1; Length 6815;

Best Local Similarity 65.8%; Pred. No. 0;

Matches 4085; Conservative 0; Mismatches 2092; Indels 27; Gaps 7;

QY 376 CTGCGAGCAAAATTTTACAGCGGCATCTTGAACATTTTACTGCTGCCGGACAGCAACGCGAT 435

Db 217 CTGTTGAAACGACATTCAGCAACGCGCTTGATCAGCTTTTACCAGTTTGAAGCGAACGCGAT 276

QY 436 CCGCTGCGTGGCGGCATCGCTCCCGAACGCTGGCGGAGGCAACGCTTTCGCTCTTTA 495

Db 277 TTATGGGGCGCTGCCATCGCGACCGCTGCGCTGGCACGAGGAAACGCTATCCGCTCACATG 336

QY 496 TTACTGCTGCTGGCAGCGCGCATATGGTGTCCGAGCTGACGCAAAATGGCGCTTCTCGAT 555

Db 337 CTGTTGTTGCTGGCAGCGCGCATCTGGCTGCAACGCCACGCCCTGCCGCTGCTCGAT 396

QY 556 CTCGCTGTGCAAGTGAATAATGTCACCGGGATCGCTGATTTCTGGATGACATTCCTCG 615

Db 397 CTCGCTGTGCGGTAGAGATGTCATCCCGCATCACTGATTTCTGGATGACATGCGCTGC 456

QY 616 ATGGATAACGGCAGATCGCTCGTGCCTACCGTGCATCGCGAATTTGGTCAAAAC 675

Db 457 ATGATGATGCGCAACTCGCTCGCGACCTCGACCATTCATTTGCCATGATGTTGGAACAT 516

QY 676 GTGCGATTTCTGCGCGCATCGCTGTAGCGCGCATTTGAAGTGAATTCGCAATGCA 735

Db 517 GTGCGATTTCTGCGCGCGTGCCTGCTGAGTAAGGATTTCCGCGTGTGCTCGTCGGCA 576

QY 736 CCGGTTTGCCTGCCATACATAAATCTGAAGGATGTCGTAACCTTCGCTGCGCTCGCTCGGC 795

Db 577 GAAGCTTTAACCGGCAACCGCCAGAGCCGACGCTGTGGCAGAAATATCCCGCAGCTCGGC 636

QY 796 CTGAGGCTTTAGTGAAGGCAATTCAGAGATCTGCACGACGCGCACGAGCGCGCAGC 855

Db 637 ATGCAAGGCTGTGTGTCAGGCGCAATTTAAGATCTCTCCGAGGTGACAGCCAGCGCAGC 696

QY 856 CCGAAGCGATGCGCATGACCAACGAACTGAAAAACCGCTGCTGTTCGCGCCACGCTG 915

Db 697 GCTGACGCCATTTGATGACCAATCACTATATAAACACGACGCTGTTCTGCGCTCCATG 756

QY 916 CAATGGCGCGATTTGCGCTGAGCTTCAACGAGGTGCGGCAAGACTTAGCTTCTTC 975

Db 757 CAGATGGCTCTATCGTGGCTGAAAGCCTCAGGTGAAGCCCGCGCAACAGCTGCAACCGTTT 816

QY 976 GCCCAGGATTTGGGCGCAGCGCTTTCACTGCTCGACGACCTCGCGAGCGTTGCAAAAC 1035

Db 817 TCGCTTAATCTTGTGTCAGGCTTTCCAGCTACTTGACAGTCTCATGACGGATGCGCCGAC 876

QY 1036 ACCGTAAGATGTGACACAGGATGAGGCAAAATCCACGCTGGTACAGATGCTCGGTGCT 1095

Db 877 ACCGTAAGATGCCATCAGGATGACGGGAAATCAACGCTGGTGAATCTGCTGGGACCA 936

QY 1096 GACGCGCGGAACGCTCGCGCGCATCTGCGCGCAGCGAGATGACACACTTGCCTGC 1155

Db 937 CAGCGGTGTAACCGGACCTCGCGGATCATCTGCGCTGCGCGCAGCGAGCATCTGTTATCG 996

QY 1156 GCCTGCCATCCGCGCATCTGCCACTCGCAATATATGACCGCGCTGTTTAATCAACAGCTA 1215

Db 997 GCCTGCCAGGACGCTTATGCCACACACCAATTTTGTTCAGGCGCTGTTGAGAAAAATC 1056

QY 1216 GCGATATTCAACTGAGCGCGCTCAGCGGTGGGCGCACTTTTGGGTTGATCGCGCGCGC 1275

Db 1057 GCTCGCGTCACTTAAGGATGCTGCA----TGAGCCACTTTTGGGCTCATTCGACCGCGCT 1111

QY 1276 TCTACAGCCATTTTCAACGCTTTCAGCGGTGAGGCTTTAGCACAAAACGCTGCTGCGCGCGCCATC 1335

Db 1112 TCTACAGCCATGTCAGCGGCTTCAGCACCTTAGCCAGGCTTAATCGCACGCGGACATC 1171



1336 GCATCATTATCCAGACGCGATGCCGCACTTTGCTTAGCGACGAAACGCATCGATT 1395  
1172 AGATTACTTTTATCCAGCAGCGAGCTTAGCGGCTACTACCGATAGCCGTATCGGCT 1231  
1396 TTGTTGCCGTCGCGCAACAGACGCACTCCCTGCGGTTGCTGCGCCGCTGTCATCGGC 1455  
1232 TTTTCCCGCTGGTTAGCCTCGCATCCGCGGCGAGTCTGGCACACACCTTGCAACTGG 1291  
1456 TGGCCTCGCGCGCGCTGCTGCTGCTTTGCGGTGATCGACGATCTCGGCTCTGACCG 1515  
1292 CGCGCATCTCTCGCGCCCTCAATGCTCAAACTGATCAATGAGATGGCCCGCAGCACCG 1351  
1516 ATATGCTGTCGCGCAACCTGCTGCGGTACTGAAGCTTTGAACATTCGATGCGGTATCG 1575  
1352 ATATGCTCTGTGTAACCTGCGCGTGTCTGAGCACCTTGGCGATTTGATGCGGTATG 1411  
1576 CCGACGAAATGGAAGCGCGCGGATTTGCTGCTGAAGCGCTGCACTGCGGCTTTGTTT 1635  
1412 TCGATCAGATGAGCCTCGCGGTGCACTGCGGCGAGAGCGCTGAACCTGCTTATGTTT 1471  
1636 CGGTGCGCTGCGCTTTGCGGTCAATGCTGAAGCGGGAATTCGCTTGGGTTGATGCCCT 1695  
1472 CGGTGCTGTGCGCTGCCACTTAACCGTGAAGCGGATTTCCGCTGGCGGTGATGCCGT 1531  
1696 TCGGTTTGCAGAGNATCAAGCGCTGAAGCGCTTTTCAGGCCAGCAGCATATCTATG 1755  
1532 TTGACTATGCCAGGACCGATCAGGCACGGAACGCTATCGCACCTGGAAGAAATCTATG 1591  
1756 ATCGCATATGCTGCTCAGCGCAGCTGATCTCAACACGCGCGGCGTTTAAATTTGA 1815  
1592 ACTGGCTGATGCACTGATCAGATCGGCTGATCGCCGCAACGCTCATGGCAATGGATTA 1651  
1816 CGGAGCGCGCGGATTTATCATGATGCTGCTGCGCTGCGCAAAATCAGCCAGATGGTC 1875  
1652 CACCACGGAAGAAATGTCATCACTGTTTTCACCGGTGCGCAGATCAACACGCTGATT 1711  
1876 CGGCTTTGATTTTCACTGACGACATGCGCGCTGCTATCAGCGCTGGGCGCACTCC 1935  
1712 CGGAATGATTTTCCCGCCAGACACTGCCCCGATTTTCAATTCGGTTGGCGCGTGC 1771  
1936 GCGCC-----CGGTTTCTCTGCGCGCTCATGCGCGCTGCGCAGCGCTGCGTCAGC 1989  
1772 GTACCACTGAACCGGCACCGCGCATCTCAGCTCGCTACTTCCGATGCGGACAGAC 1831  
1990 CGGTGCTTTATGCTCGCTGGGTACGCTGCAAGGCATCGCTTCGCGCTTTCTTTCATC 2049  
1832 CGGCAATTTTGTCTCACTCGGCACGCTCAGGGAATCGCTAGCGCTGTTTAAACCA 1891  
2050 TGGCGCAGCGTGCAGCTGCGGCTATCGCTGCTGATCGCCATTTGTGGGGATTA 2109  
1892 TCGTCCGCGCTGCCAGGATCGATGCTCAGTTATTGCTGGCGCATTCGCGTTCGTTGT 1951  
2110 ACGCGAAGCAGCATCAGCTGAGCTCGCTGGCGCGCTGGGTGACGGATTTCTGTCG 2169  
1952 CACCTTTCCAGGCGGAGAACTGGCGAGCGCCAGCATGTTGAGGTGCTGATTTTGGCG 2011  
2170 ATCAGCGCAGCCCTTACAGCAGCGCAGCTGTTTATCACTCATGCGGCTTAAACAGCG 2229  
2012 ACCAGCGCTGCACTGGCGCAGCGCGATTTGCTCATTTAGCAGCGCGCATGAATACG 2071  
2230 CGGTGAAGACTGGAATGGGTAACCGCATGCTGCGCTGCGGATTTGCTGATCAGC 2289  
2072 TGTGGATGGGATTAACCACTGACCGCTGCTGACGATTCGCGCTTTGATCAGC 2131  
2290 CGGCGTGGCGCGCATTTGATGTCATGAGCTGCTGCGCGCATCAGCTTTAGCC 2349  
2132 CGGCTGGCTGCGAGATGGTCTGGCAGGGTATCGAGCTCGCGCTCAGCTTTTACCA 2191  
2350 GTGTTTCACTCACTGAGCAGCATCTGCAACAGCTGCTGACCGCAGATCGTTAGCGCTAC 2409  
2192 CCAGCCACTCCATGGCAGCTGACCTCCAGACCTTACTGGCTGATGAAGTACGTTTACG 2251  
2410 GGATGTCAGGATTCAGGCGCAGCTGACGCGCGCGGCTTCCAGCGCTGCGCGCGACA 2469

2252 AGATGAACACGCTGCTAGCGCATTTGTCAGCGCGGCGGAACCACTTGGCGGCGACA 2311  
2470 TCGTCGAGCAGCGCTGTCGCCAGCAGCAAGTCTGCTGCGGAGCGCACTTGTATCGCAC 2529  
2312 TCGTCGAACAGCGCATGCTGACCCGCCAGCGGTCTCAC-CAGGAGACACTATGC---C 2367  
2530 GCAATACATGATGATTTTGGTCTGCTGAGCTGCGGATGCTGATTTGCTGCTGCTGCT 2589  
2368 CGGCTATGATCTGATTTCTGGTAGGTGCGGACTTGGCAACCGGCTGATCGCCCTGCGCT 2427  
2590 GGGTCAATTTGACGACCAACATGAAATGCTGCTGCTGAGAGCGATGCGCATTCGCGCAG 2649  
2428 GAGACAGAGCGGCTCTCCCTGCGCATTTCTGCTGATTTGACGCGCAACGTCGAGTGC 2487  
2650 CAATCATACCTGCTGCTTTTCAATCAGCGATCTCAGCGCGCAACAACTTCGCTGCTGCA 2709  
2488 CAATCATACCTGCTGCTTTTCAATGCGGAAGATCTCAACCGAAACGAGCATCGCTGATCG 2547  
2710 ACCGCTGATACCGTGGCTTGGTCAAGTTATCAGGTGCTTTTCTGCGCTGCGCGCA 2769  
2548 TCCGCTGCTGATGATCACTGCGCTGCTGATGAGTTCGCTTTCCCAACGAGTGCAG 2607  
2770 TCTGACGCGGATTTATTTGCTCATCGCATCAGCGATTTTGGCCGCGCATTTTACCGCGC 2829  
2608 TATGAACAGTGGCTATTTTGGTGAATCGCGAGCGCTTCGCGCAGCTCATTCGCGACAG 2667  
2830 GATGGGTGACGATCTGTGGACAAACACAGCCGTA CAACAGGTAAACACCGCAGGTGAC 2889  
2668 GTTTCGCGCGATCTGCTGCTGAATACCGGGTTCGAGCATCGCTTTACCGCGCTCAC 2727  
2890 GCTGCGGATGCGCGTGAATCTTCTGCGCAAGTGGTGTGATGCTGCGCGCTGCGAC 2949  
2728 GCTGACGATGCGCGGCTGCTGAGAGTACGAGTGAATGATGCTGCGCTGCTGAC 2787  
2950 GAGCGCAATCTGCACTGCGGTTATCAGGTGTTTCTTGAACAGAGTGCAGCTGCGCGCA 3009  
2788 CGATGCGCGCTGCGCATGGGCTTCCAGTCTGCTGCGAGAGTGCAGCTCAGCGA 2847  
3010 GCGCAACGCGCTGCGAGCAGCGCTGATGATGATGCGCACCGCTGATCAGCAAGCGGTTA 3069  
2848 GCCACATGCGCTGCGCGCGCATCATCATGATGCGACCGGTAGATCAGCAGCAGCTA 2907  
3070 TCGTTTGTCTACGCTGCGCTCAGCGCGCATCGGCTATTTGATGTAAGATACCCATTA 3129  
2908 TCGCTTTGTCTACGCTGCGCTTTTCTGACAGCACCTTCTGCTGATGTAAGATACCCACTA 2967  
3130 CGTTAACAGCGCGCTGCGCGAGAACACCGCTGCTGACGACATCGCGCATATGCCAA 3189  
2968 CATTTAATACGCTGGAAGCGATCGCGCGCTCAGAAATATCCAGCATATGCCGC 3027  
3190 TCAGCAAGCTGCGAGCTGAGTACGCTGCTGCTGAAGAGCAGCGCATATTACCGATTAC 3249  
3028 GCAGCAGGATGCGATTAGACCGGCTGCTACGCGAAGAGAGCGCTTACCCATTAC 3087  
3250 CTGAGCGGCAACATCGATCGATTCTTGGCAACAGCAGCGCGCGCAAGCTGAGCGGCT 3309  
3088 GCTGACCGCGCAGCTGCGCGCTTCTGG---CAGAACCAGATTTGCGCTGCGAGCGTTT 3144  
3310 GCGCGCGCGCTGCTGATGCGCACACCGGTTTCTTCTGCGCTGCGCGCTGCGCTGAGC 3369  
3145 GCGTGGCGGTTATTTCCATCCACCTACCGGCTATTTCTGCTGCACTGCGCGCTGCTGCG 3204  
3370 GGAGTTGCTAGCAGCGCTTGTGCCCAACCGATGCCCTCAGCTCAGCGCAACATATCGAACG 3429  
3205 TGACCGCTGGCGAGATGACAGCTTTTACCTCTGAGACCTTGCACGCCACCATTCAGCA 3264  
3430 CTTTGGCGCTGAGAGTGGCGGCAACAGCGATTTTTCGCTGCTGCTAAACCGCATGCTGTT 3489  
3265 ATTCGCACTGAGCGCTGGCAGCAACAGCGCTTCTTCCGATGCTGCTGAGCGTATGCTGTT 3324  
3490 TTTGGCGGTAGCGCAGCGCTGCGCTGATGCAACGTTTTCGCTGCGCTGCTGATGCTGCG 3549

Db 3325 TCTGGCGGCTCGGCTGACACAGCGCTGGCAGGTTATGCAACGTTTTTTACGGCCTTCCCGA 3384  
Qy 3550 CGGTTAAATTTAGCGCTTTTACGCGCGCAACTGCGCTGCGGATATAAAGCGGATTTCT 3609  
Db 3385 AGTCTGATCGCCGTTTTATGCGGAAACTCACTTTCCTGATCGGCTGGCATCTT 3444  
Qy 3610 GTGCGCAAGCCCGCGTGTCCCATCGGTGAAGCGCTGCGCGCTGTGAAATTTCTGCA 3669  
Db 3445 AAGCGCAAGCCCGCTTCCCGTTCTGCGCGGTTACAGGCTATTAATGACTCTCCACCG 3504  
Qy 3670 ACCAGGAAGAAAATGAAACGCACTTAATGATTTGGCGCAGGCTTTGCGGCTGGCG 3729  
Db 3505 TCAACAGGCGATGCAATGAATAGAACTACAGTAATTTGGCGAGGCTTTGCTGTCTGGCT 3564  
Qy 3730 CTGGCAATTCGCTGCAAGCGGCGCATACCAACCACTTACTCGAGCAGCGGACAAA 3789  
Db 3565 CTGGCAATTCGCTTCAAGGCTCAGCGGTTCCACCAGCTGCTGGAGCAGGCTGACAAG 3624  
Qy 3790 CCGGCGGACGCGCTTATGTTTGAAGACAGTGGCTTTACCTTCGATCCGCGACCCAG 3849  
Db 3625 CTGGCGGCGCGGCTTATGCTATCAGGATCAGGCTTCCAGCTTTGATCGCGGCCACG 3684  
Qy 3850 GTGATCAGGATCCAGCGCATCGAAGAGTTGTTCACTGCGCAGGAAAATGCTCAGC 3909  
Db 3685 GTAATCAGGATCCAGCGCATTTGAAGAGCTGTTTCCACCTCGCGGTAAGAGCTCTCT 3744  
Qy 3910 GATTACGTCAGCTGATCGCGGTAACGCTTCTATCGCTGTGCTGGAAGATGCAAA 3969  
Db 3745 GACTATGTCAGCTGATCGCGGTGAAGCCGTTTATCGGCTCTGCTGGAGTCCGCGAAG 3804  
Qy 3970 CAGCTTGATTAAGCAATTAATCAGCGCTGCTGGAGCAGCAGATCGCCAGTTCAATCG 4029  
Db 3805 GTGTTCAAGTTATGACAAACGATCAGCGCGCTGGAAGCGAGATGCGCGTTAATCG 3864  
Qy 4030 CAAGATGTAAGGCTATGCTCAATTTCTTCTGCTTATCAGTGAAGTATTAGAGGCT 4089  
Db 3865 CGTAGCTTGAAGGATATGCTGCTTCTGGGCTATTTCCGAGCGGTTGCTGAAGGC 3924  
Qy 4090 TATCTGAAGCTCGGACGCTGCGCTTCTGCAAGTGCCTGACATGCTGCGCTGCGCG 4149  
Db 3925 TATCTGAAGCTTGGACGCTGCGCTTCTGCTATTCGCGCAATGCTGCGCGCGGCT 3984  
Qy 4150 CAGTTGGAGCTGTGCAAGCATGCGCAGCGCTCAGCATGCTGCGCGAATTTATTCAG 4209  
Db 3985 CAGCTGGCAAACTTTCAGCGCTGCGCGCAGCGTTTACAGCAAAAGTGGCAGCTTGA 4044  
Qy 4210 GACGATCATCTGCTCAGCGCTTTCCTTCCACTCATTTGCTGCTGCGCGGTAATCTTT 4269  
Db 4045 GATGAGCATCTGCTCAGGCTTCTTTCCTTCCACTCATCTGCTGCTGCGGGAATCCGTT 4104  
Qy 4270 GCAAGCTCATCTATCTATCTTAAATTCATGCGCTGGAGCGTGAATGGGCGCTGTTT 4329  
Db 4105 GCCACTTCTCAATCTATACCTGATTCATGCGCTGGAACGCTGAATGGGCGCTGTTT 4164  
Qy 4330 CCGCGGCGGCAACGCGCGCTGCTGCAAGGCTATGCGCGCATGCTGTCAGAGCTTGGC 4389  
Db 4165 CCGCGCGGTGGACGCGCGCTGCTGCAAGGCTATGCGCGCATGCTGTCAGAGCTTGGC 4224  
Qy 4390 GCGGAGCTCTTACTGAATGCGGAAGTGAAGCAGCTGGAACCCAGCGGCAATCGCATTAGC 4449  
Db 4225 GCGGAAGTGAAGCTCAATGCCAGGTTGCGCGGCTGGAGACCCAGGAATAAGAAATACC 4284  
Qy 4450 GCGCTTCAAGTTAGGCGGAGCAGCTTTCGATGCGCGCTGCTGGCTTCCAAATCCGAC 4509  
Db 4285 GCGGTGCACTGAAGAATGGCGGCTTCTCCGACCCGCGGCTTGCCTTCCAAACGAGAT 4344  
Qy 4510 GTGCTGCATCTACGCAAACTGCTTCCGACCATCCGCTGGCAATGAAGCTGGACA 4569  
Db 4345 GTGGTTACACCTACCGGAATGCTGAGCCAGCATCCCGCTTCGAGCGCGCAGGACGA 4404  
Qy 4570 TCGCTGAAGCGTAAGCGATGAGCAACTCGCTGTTGTTGTTACTTATTTTGGCTGATCAG 4629  
Db 4405 TCACTGCAAAACAAACGATGAGCAACTCAGCTGTTTGTGATCTATTTTGGCTGATCAT 4464

Qy 4630 CCGCATGAACAGCTCGCGACCAACACCGCTCTCTTTTGGCCCGCTTATCGTAGTTGATC 4689  
Db 4465 CATCAATCAGCTGGCGACCAACACCGCTCTCTTTTGGCTCGCTTATCGTAGTTGAT 4524  
Qy 4690 GATGAGATTTTCAACAGCAGCGCAGCTGCGAGACGATTTTCACTTTTACCTGACCGGCC 4749  
Db 4525 GATGAGATTTTAAACAGATGGCTTGGCAGAGGACTTCTCGCTCTATCTGATGGGCC 4584  
Qy 4750 TGCAGAGGATTCGCTCGCTGSCACCGCCCGCTCGGCGAGCTTTTATGTTTACGCGG 4809  
Db 4585 TCGCTGACCGATCCCTCACTGCGCGCGAGGCTCGCGCAGCTACTACTGCTGCGCGCA 4644  
Qy 4810 GTGCGCGCATCTCGGACCGCTGACATCGACTGGCAACAGGAGGACCGCTTTCGCGAT 4869  
Db 4645 GTTCCGCACTCGGACCGCGATATCGACTGGCGGCTTGAAGGCTCGCGCTCGCGAT 4704  
Qy 4870 CGAATTTTGTCTTATCTGAGCAGCACTACATGCGCGGATACGTCAGCAATTAAGTACA 4929  
Db 4705 CGCATTTTGTGCTATCTGAAACAGCACTATATGCGCGGCTCGCTAGCCAGTTGCTCAG 4764  
Qy 4930 CACGAATGTTTACCGCGTTGATTTTTCGCAACAGCTGCAATGCCCATCAAGCTCTGGCG 4989  
Db 4765 CATCGCATCTTTCAGCGCTTGTATTCGCGATGAGCTGAATGCGTATCAGGCTTTCGCGC 4824  
Qy 4990 TTTTCGCTGGAGCGGATTTTGAAGCAAGCGCTGTTTCCGCGCGCATAAACCGCATGCC 5049  
Db 4825 TTCTCGTGGAGCGGATCTGACGAAGCGCTGTTTCCGCGCTCAACACCGCATAAA 4884  
Qy 5050 GATATCAGCAATCTCTATCTGCTGGTGCCTGCTAGCAATCCAGGCGCGGCTGCTCGCGC 5109  
Db 4885 AATATTAATATCTCTATCTGCTGCGTGCAGCGCATCTTGGCGCGGTTATTCAGGC 4944  
Qy 5110 GTGATCGGTTCCGCGCAAGCGCACCGCAGGCTGATCTGCGAGGATCGCGCGCATGATC 5169  
Db 4945 GTAATTTGCTCGGCGAAGCTTACCGCAGGATTTGATGCTGGAGGATCTGCGC---TTGAATA 5001  
Qy 5170 GACAGCTTTTACTTGAAGCAAGTAAGCAAGCAAGCTGCGGCTGCGTTCGAAGAGTTTCGCA 5229  
Db 5002 GTCCGCTACTGCTCGATCATGCGCTAGACACATGAGGCTGGATCGGAAAGCTTTGCA 5061  
Qy 5230 CCGCGCGCAAGCTGTTGATGACCGACCGCGCGCAGCAGCTGATGCTGATGCTGCTGCT 5289  
Db 5062 CCGCGTCAAACTGTTTGGTGCAGAACCCGACGCGCTGCTGATGCTCTACGCGCTGCT 5121  
Qy 5290 GTGCTCACTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5349  
Db 5122 GCGGTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5181  
Qy 5350 CCGTGAAGACGCGCAGGCACTATGACGATCTGCAAAATGAAACCCGCGCGCTTACA 5409  
Db 5182 CGCTGCAATCCGCGCAACAGCGCTTGGCGCAGCTGGAGATGAACAGCGTCAAGGCTATG 5241  
Qy 5410 CCGCGCGCAATGATGAAACCGCGGTTTGAAGGCTTTTGAAGGCTTTTGAAGGCTTTTGA 5469  
Db 5242 CCGGATCCAGATGATGAGCGCGCTTTGAGCGCTTTTGAAGGCTTTTGAAGGCTTTTGA 5301  
Qy 5470 AGCTGCGCAACAACTGCGCTTTCATCTGGAAGGCTTTCGCTATGATGATGATGATGATG 5529  
Db 5302 ATATTTGCTGCTTACGCTTTTTCATCTGCGCGGCTTTTTCGATGATGATGATGATGATG 5361  
Qy 5530 AACATTACGCGCTTCGATGACAGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCG 5589  
Db 5362 CACGCTACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5421  
Qy 5590 GTTTGATGATGCGCGCTTAATGCGCGTTCGCGAGGAGGCGGCTGCTGATCAGCGCTGCG 5649  
Db 5422 GCCTGATGATGCGCGCAGATTAATGCGCGCTTACGCAACAGCCAGCTGATGATGATGATG 5481  
Qy 5650 ATTTAGGAGCTGCGCTTCCAGCTTCACTAACATGCGCGGAGATTTGTAAGATGCGGAAA 5709  
Db 5482 ATCTCGTCTGGCGTTTTCAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 5541

```
QY 5710 ATGCTCGCTCTATCTGCGCAATCTCTGCTCGATCAGCGGGATTCACGCGCGATACC 5769
Db 5542 CGGAGCGCTCTATCTGCGCGCTACGCTGCTGCTGAGAGGGCTTACCGAGAAATC 5601
QY 5770 TGACTGCACGCAATCTGCGAGCGCTCGCTCTACTGCGAGCGCTTTAGTGGCGAGG 5829
Db 5602 TCGCGCATCGCAAAATCGCAGGCATTAAGCGCGCTCGCCCGCGCTGGTGGAAACGG 5661
QY 5830 CGGAACCTTATTATCACTCGGCGGATCCGGTTTACCGGTTTACCGCTCGCTCGGCT 5889
Db 5662 CAGAGCGCTTATTATCATCGGCGTGGCTGGCTGCGGTTTACCGCTCGCTTACGCT 5721
QY 5890 GGGCCATCGTACGCTCGGCGGTTTATCGGAATTCGGCTCAAAAGTTACGACCGCG 5949
Db 5722 GGGCGATTCTACCGCGCAGCAGGCTTACCGTAAATCGGTATGAAGTGGTTCAGGCGG 5781
QY 5950 GTGTGACCGCTGGGATTCACGCGCAGCGCACAGTAAAGTGAATACTGGCGCTGCTGG 6009
Db 5782 CTTTCACGCGTGGGATCAACGCCAGTCCACGACACACAGAGAACTGGCACTGCTGG 5841
QY 6010 TGAAGGGCGAGGTTTGGCGATCACTTCGCTGTGTCTCTCTGAACCGCTTCCGGCTG 6069
Db 5842 TGGCGCATCGGTACGCGGTTACTTTCGCGGTGGCGGTCAAGCTCCAGCTCCGCTG 5901
QY 6070 GTCTGCGCAGGCTCTCTGTTGATTTTACGTCGCTGACGCTGGCGAGCGTGGCTTGCAG 6129
Db 5902 ATCTCTGGCAGCGCC-----CGTTTAAAGCGCATGACGTTTCTCGCAGCGTGGCTGCAA 5955
QY 6130 CTTATTTCAGCGTGGCGGTAGAGGAAACCAACGACGACGAGCTTTCAGCGCGCGCAC 6189
Db 5956 TTTTTCACGCGCGCGCGGTAAGAAAGCAACGAGACACAGCGCTCTTTCGCTCGTAC 6015
QY 6190 CGCATGATCATCGCTGGCGCATGTATTAAGCGCTTAAAGATAGCCTTTCGCGGGATATA 6249
Db 6016 CGCATGATCATCGCTGGCGCATATAAAGCGCTGCTTAATAACCGCGCGGACATA 6075
QY 6250 GCGGAACGCGCAGGTTGATGACACGAGCGCATCTGTCACCATGAAGTAGAGCGCGCTGA 6309
Db 6076 GCGGAATGCGCAGCGCTGCTGACACAGCGCGTCAATGACAAATAATAGCAGCGCGTGA 6135
QY 6310 CGTGTCTATTTCGCGCACCAATCCATGACGCGCGCACATGCTTGCACACGACATAAT 6369
Db 6136 GAGCGTCATACCGCGCGCATCTGCAAGGCGCAGACACCGCTGCTGCCAGATAAT 6195
QY 6370 CAGCAATCGCAGTACCGCAACACACCGCATTAAGATCGTTGAGCTCAACTTACC 6429
Db 6196 GAGCAATCGCAGCGCGGACACTACGCGATAGAGTCAATTAACCTCAACGAGCC 6255
QY 6430 GCTGCGGTTATGTCGACAGATGCCAGCGCCATCCCAACCGTGCATGATGATTT 6489
Db 6256 TTTATGCGGTTATGATGTGACAGATGCCAGCGCCCAACCGCCCATGATGATGATCT 6315
QY 6490 ATGCGACGCGCGCTACCATTTCCATCACCACCGTTGCCAACAAGATAGCAGCTT 6549
Db 6316 GTGCGCGATGACGCGGTTATCTCCATCCGATAACAGTGACGACGAGCAATCAGGCGCAT 6375
QY 6550 CCATAACGACGATGTTTCTGCTCC 6573
Db 6376 CCATAACGACATAGTTCTTCC 6399
```

```
RESULT 2
ERWCT 6918 bp DNA linear BCT 24-AUG-2002
LOCUS Erwinia uredovora DNA carotenoid biosynthesis gene cluster (crtE,
DEFINITION crtX, crtY, crtI, crtB, crtZ), complete cds.
ACCESSION D90087
VERSION D90087.2 GI:22474502
KEYWORDS
SOURCE Pantoea ananatis
ORGANISM Pantoea ananatis
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Pantoea.
```

```
REFERENCE
AUTHORS Misawa,N., Nakagawa,M., Kobayashi,K., Yamano,S., Izawa,Y.,
TITLE Elucidation of the Erwinia uredovora carotenoid biosynthetic
JOURNAL pathway by functional analysis of gene products expressed in
MEDLINE Escherichia coli.
PUBMED J. Bacteriol. 172 (12), 6704-6712 (1990)
COMMENT 91072214
2254247
These data kindly submitted in computer readable form by: Norihiko
Misawa Biotechnology Institute
3-75-1, Heita, Kamaishi-shi,
Iwate 026-0001
Japan
e-mail: norihiko.misawa@bio.jp
Phone: 81-193-26-6581
Fax: 81-193-26-6584.
FEATURES
source
1..6918
/organism="Pantoea ananatis"
/mol_type="genomic DNA"
/db_xref="taxon:553"
/note="carotenoid biosynthesis genes"
225..1133
/gene="crtE"
225..1133
/gene="crtE"
/codon_start=1
/transl_table=11
/db_xref="GI:216682"
/translation="MTVCAGKHVHLTRDAEQLADIDRRLLDQLLPVEGERDVVGAAM
REGALAPKRIKRMELLTARDLGCASHGDLDLACAVEMVHAASLIIDMPCMDDA
KLRRGRPIHSHYGHVAILLAVALLSKAFGVADAGLTPLAKRANVSELNAILNQ
GLVQPFKDLSEGDKPRSAEAILMTNHFKTSTFLCASMOMAIKANASGEARDCLHRF
SLDGAQFQLLDLDLTDGTDKDSNQDAGSKTLVNLGLPRAVEERLRHQHQLASEHL
SAACQGHGATQHFIQWFDKLAAS"
1143..2438
/gene="crtX"
1143..2438
/gene="crtX"
/codon_start=1
/transl_table=11
/db_xref="GI:216683"
/translation="MSHFAATAPPYSHVRALQNLAEVLVARGHRTFTIQQYDIKHLI
DSETIGFSGVTDSPPGALTRVLHAAHPGLGSMKLINEMARTDMLCRELPQAFN
DLAVDGVTDQMEPAGALVAELGLPFI SVACALPLNRPDMPPLAVMPFEYGTSDAAR
ERYASEKIXDMLMRHDRVIAEHSRMLGAPROKLCFSLAQI SOLVPELDFPRK
ALPACFHAVGLPRLTHAPSTSSRYFTSEKPRI PASLGLTQGHRYGLFKTVKACEE
IDGQLLHACGLTDSQCEELARSHTQVDFADQSAALSQAQLTHGGMNTVLDAI
NRTPLALPLAFDQPGVASRIYHIGIGRASRFTTSHALARQMRSLLTNVDFOQRMA
KIQTALRLAGGTMAAADIIIEQVMCTGQPVLSGSYATAL"
2422..3570
/gene="crtY"
2422..3570
/gene="crtY"
/codon_start=1
/transl_table=11
/db_xref="GI:216684"
/translation="MQPHYDLILVGLANGLIARLQOQOQPMRILLIDAAPOAGGN
HTWSPHDDLTESQHRWIAPLVHWHVDPYQVRFPTRRRKLNSGYCISIQSRAEVLQR
QFGPHLWMDTVAEAVNASSVRLKGVIGARVIDGRVANSALSVGQATIGQWR
LSHPHGLSSPIIMDATVQONGRYFVLSPLSPTRLLIEDTHYIDNATLDPICARQNI
CDYAAQOQQLQTLREBQGLPITLSGNADAFWQORPLACSLGRLGPHPTTGYSLP
LAVAVADRLSALDVFTSIIHAI THFARERWQCGFFRMLNRLFLAGPDSRWVRW
ORFYGLPEDLITAFYAGKLTLDRLRLISLGRKPPVFLAALQAIMTTHR"
3582..5060
/gene="crtI"
3582..5060
```



1900 AACTGCCGCTGCTATCAGCGGTGGGCGCACTCGCGCCCGCGTTTCTCTCGCGCGC 1959  
1998 CGTTACCGCTTGTTTTATCGCGTGGCGCTCTCGCGAAACGACGCGCTCAACGT 1857  
1960 TCCATGGCGCGCTG---GCCAGCGCTGCGTACGCGGTGGTTATGCTCGCTGGGTACGC 2016  
1858 CTTTCATCCGCTTATTTTACATCTCAGAAACCCCGGATTTTCGCTCGCTGGCACGC 1917  
2017 TCGAAGGCCATCGCTTCCGCTGTTTCTGCATCTGGCGAGCGGTGCGGCCAGCTGCGGC 2076  
1918 TTCAAGGACACCGCTTATGGCTGTTTAAACAGATAGTGAAGCGCTGTGAAGAAATGACG 1977  
2077 TATCGCTGTGATCGCCATTTGGGGATTAACGCCGAAACAGACGATCAGCTGGAGC 2136  
1978 GTACGCTCTGTTAGCCCACTGTGCTTACCGACTCTCAGTGTGAAGAGCTGGCGC 2037  
2137 TCGCTGGCGCGGTGGGTGACGATTTTCGTGCATCAGCGCGCAGCGCTTACAGCAGCGC 2196  
2038 GAAGCGCTCATACAGGTGGTGGATTTTCCGATCAGTCAGCGCGGCTGTCTCAGGCGC 2097  
2197 AGCTGTTTATCACTCATGCGCGGTTAAACAGCGCGCTGGAAGCACTTGGAAATCGGTACGC 2256  
2098 AGCTGCGGATCACCAAGCGCGCATGATACGCTACTGGACGCGATTAAATACCGGAGCG 2157  
2257 CGATGCTGCGCTGCGGATGCTTTGATCAGCGCGCGCTGGCGCGCATGAGTGCG 2316  
2158 CCGCTTTTACGCTTCGCTGGCGCTTTGATCAGCGCGCGCTCGCTCAGCGATCGTTATC 2217  
2317 ATGACGCTGCTGCGCGCATCAGCGCTTTAGCGCTTTCATCAACTGGAGCAGCATTCGC 2376  
2218 AGCGCATCGGCAAGCGTGTTCGCGCTTTACACGCGCATGCTTTGGCTGTGATGCG 2277  
2377 AACAGCTGCTGACCGAAGATGTTAGCGCTACGGATGTCAGCGATTACGGCGCAGCTGC 2436  
2278 GTTCATTGCTGACCAAGTCGATTTACAGCGCATGGCGAAATCCAGACAGCGCTTC 2337  
2437 AGCGCGCAGCGGTGCGAGCTGCGCGCGCATGCTCAGCAGCGCGTGTGCGAGCAGC 2496  
2338 GTTTGCGCAGGGGCAACATGGCGCGCTGCGCGATCATCATTGAGCAGGTTATGTG-CACCGGT 2396  
2497 AAGTGTGCTGCGGAGGCGACTGATGCGCAGCAATACGATGTGATTTGCTCGGTGC 2556  
2397 CAGCCTGCTTAAAGTGGAGCGGCTATGCAACCGCATTAATGATCTGATTTCTGTGGGCG 2456  
2557 TGGACTGGCGAATGGCTTGATTTGCGTGTGCTGCGTCAATTCAGCCCAACACTGAAATG 2616  
2457 TGGACTCGGAATGGCTTATCGCCCTGCTCTTACAGCAGCAGCAACCTGATATGCTAT 2516  
2617 CCGTGTGCTGGAGAGGATGCGATCCGCGAGGCAATCATACTGCTGTTTTCATCAGCAG 2676  
2517 TTTGCTTATCGACCGCGCACCCAGCGGCGGGAATCATAGCTGGTCAATTTTCAACACGA 2576  
2677 CGATCTCAGCGCGGAACAACTTCGCTGCTGCAACCGCTGATTACCGCTGTTGTTGTCAGG 2736  
2577 TGAATTTGACTGAGGCAACATCGTTGGATAGCTCCGCTGGTGGTTCATCTAGCCCGGA 2636  
2737 TTATCAGGTGCGTTTCTCGCTGCGCGCAATCTGGAACGGGATTAATTTTCCATCGC 2796  
2637 CTATCAGGTACGCTTTCCACACGCGCTGTAAGCTGGAACAGCGGCTACTTTTGTATTAC 2696  
2797 ATCAGCGATTTTTCGCCGCTTTCAGCGGCGATGGGTGAGGATCTGTGTGACAAACAC 2856  
2697 TTTCTCAGCGTTCGCTGAGGTTTTACAGCGACAGTTTGGCCCGCACTTGTGTGATGGATAC 2756  
2857 AGCGGTACAAAGGTTAAACCCAGCGAGTGAAGCTGCGCGGATGGCGTGAATCTGCTGC 2916  
2757 CCGGCTCGCAGAGGTTAATGCGGAATCTGTTCCGTTGAAAGGCTCAGGTTATCGGTGC 2816  
2917 GCAAGTGTGATTTGAGTGTGCGGCGCTGACGCGGACCGCACATCTTCAGCTGGGTTATCA 2976  
2817 CCGCGCGGTGATTTGACGGCGGGGTTATGGCGCAATTCAGCAGTGGGCTTCCA 2876  
2977 GGTGTTTCTTGGACAAGAGTGGCAGCTGGCGGACGCGCGCTGCGCAGCGCGATCTC 3036

2877 GCGCTTATTTGCCAGGAATGGCATTTAGCCACCAGATGGTTTATCTCTCTCCCATTTAT 2936  
3037 GATGATGCGACCGCTCGATCAGCAAGCGGTTTATGTTTGTCTACACGCTCGCGCTCAG 3096  
2937 CATGATGCCAGGTCGATCAGCAAAATGGTTATCGCTTCGTGTACAGCTCGCGCTCTC 2996  
3097 CGCGCATCGGCTTATTTGATGAAGATACCATTTACGTTAACAGCGCGCTGGCGAGAA 3156  
2997 GCGGACCAAGATTTGTTAAATTTGAAGACACGCACTATATTGATAATGGACATTAAGTCTGA 3056  
3157 CACCGCTGTCAGCAGCATCGCGATATGCCAATCAGCAAGCTGAGACCTGAGTACGCT 3216  
3057 ATGCGCGGCGGCAAAATTTTGGACTATGCGCGCAACAGGGTTGGACGCTTCAGACACT 3116  
3217 GCTGCTGGAAGCAGCAGCATATTACCGATTACCTTGAGCGGCAACATCGATCGATTCTG 3276  
3117 GCTGCGAGNAGAACAGGGCGCTTACCATTTACTCTGTGCGGCAATGCCGACGCACTCTG 3176  
3277 GCAACAGCAGCGCGGCGCAAGCGTGCAGCGGCTGCGCGCGGCTGTTTTCATGCCACAC 3336  
3177 G---CAGCAGCGCGCGCTGCGCTGTAGTGGAATTAGTGGCGCTGTCTTCCATCTACCA 3233  
3337 CGGTTACTCTTTCGCTGCGCTGCGCTGAGCGAGTTGTTAGCAGCGCTGTGCGCCAC 3396  
3234 CGGCTATTACCTGCGCGTGGCGGTTGCGTGGCGCACCGCTGAGTGCACTTGATGCTTT 3293  
3397 CGATGCCCTCAGCTCAGCCACATATCGAACGCTTTGCCCTGACGAGTGGCGCGCAACA 3456  
3294 TAGCTGCGCTCAATTCACCATGCCATTACGCAATTTGCCCGGAGCGTGGCAGCA 3353  
3457 GCGATTTTTCGCTGTCTGTAACCGCATGCTGTTTTTGGCGCGTAAAGCGCAGAGCGCTG 3516  
3354 GGGCTTTTTCGCTGCTGAATCGCATGCTGTTTTTAGCGCGACCGCGGATTACGCTG 3413  
3517 GCGCTGTATGCAACGTTTTTACCGCTCGATGCGGTTAAATAGCGCTTTTACGCGCG 3576  
3414 GCGGTTATGACGCGTTTTTATGTTTACCTGAAGATTAAATGCGCGTTTTTATCGGG 3473  
3577 GCAACTGCGCTCGCGCATAAACCGGATTCTGTGCGCAAGCGCGCTGCGCATCGG 3636  
3474 AAACTCAGCTGACCGATCGCTAGCTATTCTGAGCGGCAAGCGCTTTCGCTATT 3533  
3637 TGAAGCGCTGCGCGCTGTTGAAATTTCTGTCGAAACAGGGAAGAAAAATGAACGCACT 3696  
3534 AGCAGATTGCAAGCGCAATTTATGACACTCATCTTAAAGAGCGCACTACATGAACCACT 3593  
3697 TATGTGATTTGGCGAGGCTTTGGCGCTGCGCTGGCGATTTCGCTGCAAGCGCGGCG 3756  
3594 ACGGTAATTTGGTGAGGCTTCGCTGGCTGCGCAATTCGCTTACAAGCTCGCGG 3653  
3757 ATACCAACCACTTACTCGAGCAGCGCAAAACCGGCGGAGCGCGCTATGTGTTGAG 3816  
3654 ATCCCGCTTACTCTTGAACAAAGTATTAACCGCGGCTGGGCTTATGTCTACGAG 3713  
3817 GACAGTGGCTTACTCTGATGCCGACCCAGCGTGAATCAACGATCCGAGCGCATCGAA 3876  
3714 GATCAGGCGTTTACCTTTGATGAGCGCGCGAGCGTTATCAGGATCCGAGTGCATTGA 3773  
3877 GAGTTGTTACGCGAGGAATTCGCTCAGCGATTAGCTGAGCTGATGCGGTAAACG 3936  
3774 GAATGTTTGCCTGGCAGGAACAGTTAAAGAGTATGTGAACTGCTGCGGTTACG 3833  
3937 CCCTTCTATCGCTGCTGGGAAGATGGCAACAGCTTGAATTAACGACATTAATCAGCG 3996  
3834 CGGTTTACCGCTGTTGGAGTCAGGGAAGCTTTTAAATTAAGTAACGATCAAC 3893  
3997 CTGCTGAGCAGCAGATCCGACGTTCAATCCGCAAGATGTAGAAGCTTATCGTCAATTT 4056  
3894 CGGCTCGAAGCAGCAGATTTCAGCAGTTTAAATCCCGCGATGTGCAAGGTTTATCGTCA 3953  
4057 CTTGCTTATTCAGTGAAGTATTTAGAGGAGGTTTCTGAAACTCGGCAAGCTGCGCTT 4116



Db 3954 CTGGACTATTCAACGCGCGGTGTTTAAAGAGGCTATCTAAAGCTCGGTACTGTCCCTTTT 4013  
Qy 4117 CTGCAAGTGCATGACATGCTGCGCTCGCGCGAGTTGGGACGCTCGCAACATGCGCG 4176  
Db 4014 TTATCGTTTCAGAGACATGCTTCGCGCGACCTCAACTGGCGAAACTGCGAGCATGGAGA 4073  
Qy 4177 AGCGCTACAGCATGGTGGCGAAATTTATTTACAGACGATCATCTGCGTCAGCGGTTTTCC 4236  
Db 4074 AGCGTTTACAGTAAGTGGCCAGTTTACATCGAAGATGAACATCTGGCCAGCGTTTTCT 4133  
Qy 4237 TTCCATCATTTGCTGGTGGCGGTAATCTCTTTTGAACGTCATCGATCTATACCTTAAT 4296  
Db 4134 TTCCATCTGCTGTTGGTGGCGGCAATCCCTTCGCGACCTCATCTCATTTATACGTTGATA 4193  
Qy 4297 CATGCGTGGAGCGTAATGGGCGCTGTGTTTCCGCGCGCGGCAACCGCGCGCTGTG 4356  
Db 4194 CAGCGCTGGAGCGTAGTGGGCGCTGTGTTTCCGCGTGGCGGCAACCGCGCGCATAGTT 4253  
Qy 4357 CAGGCGATGGCGGACTGTTCAGGACTTTGGCGCGGAGCTGTTACTGAATGCCGAAGTG 4416  
Db 4254 CAGGCGATGATAAGCTGTTTCAGGATCTGGGTGGGAAGTCTGTATAACGCCAGAGTC 4313  
Qy 4417 AGCCAGTGGAAACAGCGGCGAATCGCATAGCGGTTTCAAGTATAGAGGCGGACGACGC 4476  
Db 4314 AGCCATATGGAACAGCAGGAAACAGGATTTGAAGCGGTGCATTTAGAGGACGGTCGACG 4373  
Qy 4477 TTCGATGCGCGCTGTGCGCTCCAAATGCCAGCGTGTGCATCTACGACAAATCTGTT 4536  
Db 4374 TTCTGACGCAAGCGTTCGCTCAATGCAATGAGATGTGTTTACCTATTCGCGACCTGTTA 4433  
Qy 4537 GCGCACCATTCGCTGGCAATGAAGCTGGACATCGCTGAACGCTGAAGCGCATGAGCAAC 4596  
Db 4434 AGCCAGCACCTCGCGCGGTAAAGCAGTCCAACTGCAGACTAAGCGCATGAGTAAC 4493  
Qy 4597 TCGCTGTTTGTACTCTATTTGGCTGAAATAGCGCGATGAAACAGCTGCGCACACACC 4656  
Db 4494 TCTCTGTTGTCTCTATTTGTTGAATCAACCATCATGATCATGATCAGCTCGCGCATCACAG 4553  
Qy 4657 GTCGTTTTGGCGCGTTATCGTCAAGTTGATCGATGAGATTTCAACAGCAGCCAGCTG 4716  
Db 4554 GTTGTGTTGGCGCGGTTTACCGAGCTGATGACGAATTTTAATCATGATGGCTTC 4613  
Qy 4717 GCAGACGATTTTCTACTTACCTGACGCGCTCGCAGCAGCATCCGTCGCTGGCACCG 4776  
Db 4614 GCAGAGGACTTCTCACTTATCTGACGCGCTCTGTGTCAGGATTCGTCACCTGGCGCT 4673  
Qy 4777 CCGGCTGGGACGCTTTTATGTTTATGTTAGCGCGGTGCGCATCTCGGACACGCTCAATC 4836  
Db 4674 GAAGGTTGCGGAGTTACTATGTGTTGGCGCGGTGCGCATTTAGGCAACCGCGAACCTC 4733  
Qy 4837 GACTGGCAACAGGAAGGACCGGCTTTCGCGATCGAATTTTTCGCTTATCTGGAGCAGCAC 4896  
Db 4734 GACTGGAAGTTGAGGGGCAAACTACCGGACCGTATTTTTCGCTACCTTGAGCAGCAT 4793  
Qy 4897 TACATCCGGGATTAAGTCAGCAATTTAGTGACACACAGAAATTTTACGCGCTTTGATTTT 4956  
Db 4794 TACATGCTGCTTACGAGTCAAGTGGTCAGCAGCAGCATGTTTACGCGCTTTGATTTT 4853  
Qy 4957 CCGGACAGCTCATGCCCATCAGGCTCGGCTGCGGTTTTGCTGGAGCGAATTTGACGCA 5016  
Db 4854 CCGGACCAAGCTTAATGCTTATCATGGCTCAGCTTTTCTGTGGAGCGGCTTCTTACCCAG 4913  
Qy 5017 AGCGCTGTTCCGCGCGCATAAACCGCATGCCGATATCAGCAATCTCTATCTGCTGGGT 5076  
Db 4914 AGCGCTGTTTCGCGCGCATAAACCGCATAAACCAATTAATCTCTACCTGCTGGC 4973  
Qy 5077 GCGGTAACGATCCAGCGCGGCGTGGCGGCTGATCGGTTTCGCGCAAGGCCACCGCC 5136  
Db 4974 GCAGGACGATCCGCGCGAGGCAATCTGCGCTCATCGGCTCGGCAAAAGCGACGCA 5033  
Qy 5137 AGGCTGATCGGAGGATCGCGCGGATGAATGCAATCGACGCTTTACTTGGCAAGTAAACGC 5196  
Db 5034 GGTGTTGATGCTGGAGATC---TGATATGAATAATCCGCTGTTACTCAATCATCGCGTGC 5090

Qy 5197 AAACCATGCGGTGGCTCGAAGAGTTTCGCCACCGCGCCAAAGCTGTTTGTATGACACCA 5256  
Db 5091 AAACCATGCGGTGGCTCGAAGAGTTTCGCCACCGCTCAAAGTATTTTGTATGACAAA 5150  
Qy 5257 CCGCGCGCAGCAGCTGATGCTGTATGCTGTGTGCTCACTGCGATGATGATGATGATG 5316  
Db 5151 CCGCGCGCAGCAGCTGATGCTGTACGCTGTGTGCGGCAATTTGACGATGTTTATTGACG 5210  
Qy 5317 GGCAAAACCTGGGCGAAGCGCGCAGCATGCCGTGGAAGACCGCGCAGGACGATGTC 5376  
Db 5211 ATCAGACCTGGGCTTTTCAAGCCCGGCGAGCTGCTTACAAACGCGCAACAACTGTA 5270  
Qy 5377 AGCATCTCAAAATTCAAACCCGCGCTTACAGCGCGCGCACATGATGATGAACCGGCT 5436  
Db 5271 TGCAATTTGAGATGAACCGCGCGAGCTTATGCAAGATGCGATGACAGAACCGGCT 5330  
Qy 5437 TTAGGCGCTTTTCAAGAGTGGCGATCAATTCACAGCTCGCGCAACAACTGGCGTTGATC 5496  
Db 5331 TTGCGCTTTTCAAGAGTGGCTATGCTGCTCATGATATCGCCCGGCTTACGCGTTGATC 5390  
Qy 5497 ATCTGGAAGGCTTCGCTATGATGCAACGCAACAACTTACGAGCTTCGATGACACGC 5556  
Db 5391 ATCTGGAAGGCTTCGCTATGATGCAACGCAACAACTTACGAGCTTCGATGATACGC 5450  
Qy 5557 TGCCTTACTGCTTATCACTGCGCGCTGTGCTGCTTGTATGATGCGCGCGCTAAATGGGCG 5616  
Db 5451 TGCCTTATTTGCTATCACTGCGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5510  
Qy 5617 TGCAGCAAGAGCGGTGCTGATCAGCTGCGATTTAGGACTGCGCTTCAGCTCACTA 5676  
Db 5511 TGCAGCAAGAGCGGTGCTGATCAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5570  
Qy 5677 ACATTTGCGCGGACATTTAGAGATGCGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 5736  
Db 5571 ATATTTGCTGCGGATATTTGAGACGATGCGCATGCGCGCTGCTGCTGCTGCTGCTGCT 5630  
Qy 5737 GCTCGATCAGCGGGAATTAACGCGCGATACGCTGACTGACCGCAACATCTGTCAGCGC 5796  
Db 5631 GCTCGGAGCATGAGGCTCTGAACAAAGAGATTTATGCGGCACTGAAACCGCTCAGGCGC 5690  
Qy 5797 TCGCCTACTGCGAGCGGTTTATGCGGAGGCGGAAACCTTATTTATCACTCGGCGCAT 5856  
Db 5691 TGAGCGGTATCGCGCTGCTTGTGTCAGGAGCAGAACTTACTTATTTGCTGCGACAG 5750  
Qy 5857 CCGGTTTACCGGTTTACGCTGCGCTGCGGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 5916  
Db 5751 CCGGCTGCGAGGTTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCTGCT 5810  
Qy 5917 ATCGCAAAATTTGCGCTCAAAGTTTACAGCACGCGGCTGTCAGCGCTTGGGATTCACGCGCAGC 5976  
Db 5811 ACCGAAATAGGTGTCAAAGTTTGAACAGGCGGTCAGCAAGCTTGGGATCAGCGGCACT 5870  
Qy 5977 GCACAGTAAAGGTGAAAACTTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6036  
Db 5871 CAACACACGCGCGGAAATTAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5930  
Qy 6037 CCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6096  
Db 5931 CCGGATGCGGCTCATCTCTCCCGCTGCGCATCTCTGCGAGCGCGCTGCTGCTGCTGCT 5984  
Qy 6097 AGCTCCGTGAGCTGCGCGCAGCGTGGCTTTCAGCTTATTTTCAGCGGTGCGCGCTAGAGGAA 6156  
Db 5985 AGCGCATGCTCTTCCCGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6044  
Qy 6157 ACCAAACGACGAGCGCTTTCAGCGCGCGCAACCGCATGATGATGCTGCGGCTGCTGCTGCT 6216  
Db 6045 GCGAAAGAAACCAACCTTCTTTGCGCTGACGCGCTGATGATGATGATGATGATGATGATGAT 6104  
Qy 6217 TAAGCGCTTAAGATGACCTTTGCGCGGATATAGCGGACGCGCGCTGCTGCTGCTGCTGCT 6276  
Db 6105 CAACGTTTGAAGTACCTTTCGCTGGAATATAGCGGAATAGCGGAACGTTTGTATGACCCAG 6164

```
QY 5277 GCCATGTCACATGAAGTAGAGCGCGGTACGTCGTCATTCGGGACGATCCACTG 6336
Db 6165 CCGCGTGTGACACATAAAATAGTAATCCATACGCGGTACATCGGCCAATCCACTG 6224
QY 6337 CAGCGGCCACATGCTGTTGACACACGACATATAATCAGCAATCGCCAGTACCGCAAAAC 6396
Db 6225 GAGCGGCCACATTCCTGTACTGCCAGATATAATCAGCAGATCGATATGAGCAAAAC 6284
QY 6397 CACCGGCAATGAAGTCTGAGCTCAAACTTA CCGCTGTGCGGTTTCATGTGCGACAGATG 6456
Db 6285 CACGGCATAAAGATCGTTAACTTCAACGACGCTTTACGCGGTTTCATGATGTGAAGATG 6344
QY 6457 CACGCCCATCCCAACCGTGCATGATGATTTATTCGACAGCGCGCTACCATTTCCAT 6516
Db 6345 CCATCCCAACCCAGCGGTGCATGATGATTTGTGTGCGAGTACGATCCTTTCCAT 6404
QY 6517 CACCAACCGGTTGCAACAAAGATAAGCACTGTTCCATACCAAGAGCATTTGTTCTGTCATT 6576
Db 6405 GCCATCAGGTTAAAGAAACGATCAGGCAATCCAAATCCCAACATATTTCTCCGGT 6464
QY 6577 TGTGAAAGGGAAGTA 6593
Db 6465 AGAGACGTCGTGCAGCA 6481

RESULT 3
112910
LOCUS I12910 6918 bp DNA linear PAT 26-JUL-1995
DEFINITION Sequence 13 from patent US 5429939.
ACCESSION I12910
VERSION I12910.1 GI:910887
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 6918)
AUTHORS Misawa,N., Kobayashi,K., Nakamura,K. and Yamano,S.
TITLE DNA sequences useful for the synthesis of carotenoids
JOURNAL Patent: US 5429939-A 13 04-JUL-1995;
FEATURES
Location/Qualifiers
1..6918
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 37.5%; Score 2624.6; DB 6; Length 6918;
Best Local Similarity 64.7%; Pred. No. 0;
Matches 4007; Conservative 0; Mismatches 2169; Indels 21; Gaps 6;

QY 400 CTTGAACATTTACTGCTCCGGACAGCAAGGATCGCGTGCCTGCGCGATGCGTGC 459
Db 303 CTTGATCAGTTATTGCCGTGGAGGAGAAACGGGATGTTGTGGGTGCGCGATGCGTGA 362
QY 460 GGAACGCTGCGCAGGCAAGATTCGTCCTTTTACTGCTGCTGCGCAGCGCGAT 519
Db 363 GGTGCGCTGGCACCAGGAAACGATATTCGCCCATGTTGCTGTGCTGACCCCGCGCAT 422
QY 520 ATGGGTGCGAGTGAACAAATGCGGTTCTCGATCTCGCTGCGAGTGAATGTTG 579
Db 423 CTGGGTGCGCTGTCAGCATACGGAATTACTGGATTTGGCTGTGCGGTGGAATGTC 482
QY 580 CACCGGCAATCGCTGATCTGATGACATTCCTCGATGGAATGGAATGGAATGCGTGT 639
Db 483 CACGCGGCTTCGCTGATCCTTGAAGATATGCGCTGATGGAATGGAATGGAATGCGTGT 542
QY 640 GGTGCGCTACCGTGCATCGCAATTTGTTGAAAGTGGCGATTCGCGCGCATCGCG 699
Db 543 GGACGCGCTACCATTCATTTCTATACGAGAGCATGTGGCAATACTGCGCGGTGTC 602
QY 700 CTGCTTAGCGCGGCAATTTGAAGTGAATTCGATTCACCGGTTTGGCTGCCATACATAA 759
Db 603 TTGCTGATGAAGCCCTTGGCGTAATTTGGCGATGAGATGGCTTCAGCGCGCTGGCAAA 662
```

```
QY 760 TCTGAAGCGATTTGCTGAATCTCCGCTGCGCTGCGGCTTAGTGCAAGGGCAA 819
Db 663 AATCGGGGGTTTCTGAATCTGTCAACCGCAATCGGCAATGCAAGGATTTGGTTCA 722
QY 820 TTCCAGGATCTGACAGACGCGACAGCGAGCGCGAGCGCGGAAAGCGATCGCCATGACCAAC 879
Db 723 TTCAAGGATCTGCTGAAGGGGATGAAGCGCGAGCGGCTGAAGCTATTTTGATGAGCAAT 782
QY 880 GAATGAAACCAAGCGTCTGTTTCGCGCCAGCTGCAAAATGCGCGCGGATTCGCGCTGAC 939
Db 783 CACTTTAAACCCAGCAGCTGTTTGTGCTCATGACAGATGGCCTCGATTTGTCGAAT 842
QY 940 GCTTCACCGAGTTCGCGCAAGACTTAGCTTCTTCGCCAGGATTTGGCGCGAGCGTTT 999
Db 843 GCCTCCAGCGAAGCGGATGTCCTGATCGCTTTTTCATTTGATCTTTGGTCAGGCAATTT 902
QY 1000 CAACTGCTCGACAGCTCGCGGACGCTTGAACAAACACACCGGTAAAGATGTGCAACAGGAT 1059
Db 903 CAACTGCTGACGATTTGACCGATGCGCATGCGGACACCGGTAAAGATGAGCAATCAGGAC 962
QY 1060 CAGGCAAAATCCAGCTGCTGATGATGCTGCGTGTGAGCGCGCGGAAAGCTTCGCTGCGC 1119
Db 963 GCGGTAATTCGACGCTGCTCAATCTGTAGCGCGGAGCGGCTTGAAGAACTGTGAGA 1022
QY 1120 GATCAGCTGCGCAGCGCAGATGCAACCTTGCCTGCGCTGCGATCGCGGATCGCCACT 1179
Db 1023 CAACTCTTCAGCTTTCGCGAGTGCATCTCTTCGCGGCTGCGCAACACGCGGACGCGACT 1082
QY 1180 GCGCAATATATGACGCGCTGTTTAAATCAACAGCTAGGATATTAATCACTAGCGCGGCTC 1239
Db 1083 CAACTATTTATTCAGCGCTGTTTGAACAAACCTGCTGCGCTCAGTTAAGATGCTGC 1142
QY 1240 AGCGGTTGGGCACTTTTGCCTGATGCGCGCGCTCTTACAGGCACTTTTCAAGGCTTGC 1299
Db 1143 A-----TGAGCCATTTTCGCGCGATCGCACGCGCTTTTACAGCCATGTTTCGCGCATTA 1197
QY 1300 AGCGGTTAGCAAAAGCTGCTGCGCGCGCGGCGCATCGCATCATTCATCATCGACGCGG 1359
Db 1198 AGAATCTCGCTCAGGAACCTGCTGCGCGCGGCTCATCGGCTGACCTTTTATTCAGCAATACG 1257
QY 1360 ATGCGCGCACTTTGCTTAGCGACGAAACGATCGATTTTGTGCGCTGCGGCAACAGACGC 1419
Db 1258 ATATTAAACATTTGATCGATAGCGAAACATTTGATTTTCATTTCCGTGCGGACAGACGC 1317
QY 1420 ATCTGCGCGGTTGCTGCGCGCGCTGTTGCTGATGCGCTGCGCTGCGCGCGGCTGCTGCG 1479
Db 1318 ATCCCCCGCGCGCTTAACGCGCGTGTACACCTGCGGCTCATCTCTGCGCGCGCTCAA 1377
QY 1480 TGTTCGCGTGTATCGACGATCTGCGCTCTGACCGGATATGCTGCGCGGAACTGCGCTG 1539
Db 1378 TGCTGAAGCTCATCAATGAAATGGCGCGCACCCACGATATGCTGCGCGGAACTCCCC 1437
QY 1540 CGGTACTGAAAGCATTTGAACATCGATGCGGATGCGGCAAGATGGAAGCGCGCGGCG 1599
Db 1438 AGGCATTTAAGATCTGCGCGCTGATGCGGCTCATTTGTTGATCAATGGAACCGCGGCGG 1497
QY 1600 GATTGCTGCTGAAGCGCTGCAATGCGCTTTGTTTCGCTGCGCTGCGCTTTCGCGGTTCA 1659
Db 1498 CGCTGTTGCTGAAGCACTGGGACTGCGCTTTATCTGCTGCGCTGCGGCTGCGCTCTCA 1557
QY 1660 ATGCTGAAGCGCGGATTCGCTTGGGTGATGCGCTTTCGCTTTTTCGACAGGATGACAAAG 1719
Db 1558 ATCTGTGAACCGGATATGCGCTTGGCGGTTATGCTTTTTCGAAATACGGGACACGCGCG 1617
QY 1720 CGCTGAAACGTTTTCAGGCGCAGCAGATATCATGATCGCATATGCTGCTGCTGCGGCG 1779
Db 1618 CTCGGAACGTTATCGCGCGAGTGAAATTTTATGATGCGGCTATGCTGCTGCTGATGACC 1677
QY 1780 AGTGATCTTCAAAACACGCGCGGCGGTTTAAATTCAGCGAGCGCGCGGATTTACATCACT 1839
Db 1678 GTGTCTATGCGGAAACACAGCCACAGATGCGGCTTAGCGCGCGCGGCAAAAGCTTCACTCA 1737
QY 1840 GCCTGTCGCGCTGCGCAAAATCAGCAGATGCTGCGCGGCTTTGATTTTTCACGCTGACG 1899
```

1738 GTTATTTTCGCACTGGCGCAATCAGCAGCTTGTCTCTGAACTGGATTTTCGCCGCAAG 1797  
1900 AACTGCCCGCTGCTATACGCGCTGGGCGCACTCGCGCCCGCGTTCCTCTGCGCGC 1959  
1798 CGTTACCGGCTGTTTTCATCGCGCTGGGCGCTCTGCGGAAAGCAACGACCGTCAACGT 1857  
1960 TCATGCGCCGCTG- --GCCAGCGCTGCGTCAGCGCGTGGTGTATGCTCGCTGGGTACGC 2016  
1858 CTTTCATCCCGTTTATTTTATCATCTCAGAAAAACCCCGGATTTTCGCTCGCTGGGACGC 1917  
2017 TGCAGGCCATCGCTTCGCGCTGTTTCTGCACTGCGCGAGGGGTGCCCGAGCTCGCGC 2076  
1918 TTCAGGGAACCGGTTATGGGCTGTTTAAACAGCATAGTGAAGCGCTGTGAAGAAATTTGACG 1977  
2077 TATCGCTGCTGATTCGCGCCATTTGGGGGATTTAAACGCCGAAACAGACGCATCAGCTGGAGC 2136  
1978 GTACGCTCTGTTAGCCCACTGTGGTCTTTACGGACTCTCAGTGTGAAGAGCTTGGCGC 2037  
2137 TCGCTGGCGGGGTGGGTGACCGGATTTCTGATCAGCGCGCAGCCCTACAGCAGCGCGC 2196  
2038 GAAGCGCTCATACAGAGTGGTGGATTTTTCGCGATCAGTCAGCGCGCTGTCTCAGGCGC 2097  
2197 AGCTGTTTATCATCATGCGCGGTTAAACAGCGCGCTGGAAGCATGGAATCGGTTACGC 2256  
2098 AGCTGGCGATCACCCACGCGCGCATGAATACGGTATCTGGACGCGATTAATTAACCGAGCGC 2157  
2257 CGATGCTGCGCTGCGCGATGCTTTTGTATCAGCGCGCGTGGCGGCGCATTTGAGTGGC 2316  
2158 CCGTTTATGCGCTTCGCTGGCTTTGATCAGCGCGCGCTCGCTCAGCATCGTTTATC 2217  
2317 ATGAGTGGTGGCGCGCATCACGCTTTAGCGTGTTCATCAACTGGAGCAGCATCTGC 2376  
2218 ACGGATCGCAAGCGTCTTCCCGCTTTTACCACGAGCATGCTTTGGCTGCTCAGATGC 2277  
2377 AACAGCTGTGACCGACGATCGTTAGCGCTAGGATGTCAGGATTCAGGGCGAGCTGC 2436  
2278 GTTCATGTGTGACCAACGCTGAGCTTTCAGCAGCGATGCGGAAAAATCCAGACAGCGCTTC 2337  
2437 AGCGCGCAGCGGTTGCCAGCGTGCGCGACATCGTCAGCAGCGCTGTCAGAGCAGC 2496  
2338 GTTTGGCAGGGGCACTAGCGCGCTGCGGATCATTTGAGCAGGTTATGTG- CACCGT 2396  
2497 AAGTGTGTGCGGAGGCGAAGCTGATGCGCAACGAAATACGATGATGATTTGGTTCGCTGC 2556  
2397 CAGCCTGTCTTAAGTGGGAGCGGCTATGCAACCGCATTTATGATCTGATCTCGTGGGCGC 2456  
2557 TGGACTGGGAATGCTGATTTGCGCTGCTGCTGCTCAATTTGACGACCAACTGAATG 2616  
2457 TGGACTCGGAATGGCTTTATCGCCCTGCGTCTT CAGCAGCAGCAACCTGATATCGTAT 2516  
2617 CCGTGTGCTGAGAGCGATGCGCATCCGGCAGGCAATCATCTGCTGCTTTCATCAGCAG 2676  
2517 TTTGCTTATCGAGCGCGCACCCAGCGGGCGGGATCATAGTGGTCAATTCACACGA 2576  
2677 CGATCTCAGCGCGCAACAACTTGTGCTGCTGCAACCGGTGATTAACGCTGCTGCTCAGG 2736  
2577 TGATTTGACTGAGAGCAACATCTGTTGATAGCTCCGCTGGTGTTCATCACTGSCCCGA 2636  
2737 TTATCAGGTGCTTTTCTGCGCTGCGCGCAATCTGGACGGGATTTATGTTCCATCGC 2796  
2637 CTATCAGGTACGCTTTCCACACGCGCTGTAAGCTGAAACAGCGGCTACTTTTGTATTAC 2696  
2797 ATCAGCGGATTTTGGCCGCACTTTTACGCGCGATGGGTGACGATCTGTGGACAAACAC 2856  
2697 TTCTCAGCGTTTCTGAGGTTTACAGCGACAGTTTGGCCCGCACTTGTGGATGGATAC 2756  
2857 AGCCGTAACAAGGTAAACCAACGAGGTGACGCTGGCGGATGCCGTGAACCTTGTGC 2916  
2757 CCGGCTCGCAGAGGTTAATGCGGAATCTGCTCGGTTGAAAAAGGCTCAGGTTATCGGTGC 2816  
2917 GCAAGTGTGATTTGATGTTGCGGCGCTGCAGCCGACGCGCATCTGAGCTGGGTTATCA 2976

2817 CCGCGCGTGTATTGACGGCGGGGTTATGTCGGGAAATTCAGCACTGAGCGTGGGCTTCCA 2876  
2977 GGTGTTTCTTGACAAAGATGCGACTGGCGCAGCGCGCACGCCCTGCAAGCAGCCGATCCT 3036  
2877 GCGGTTTATTTGCCAGGAATGGCGATTGAGCCACCAGCATGGTTTATCGTCTCCCATTTAT 2936  
3037 GATGATGCCACCGTTCGATCAGCAAGCGGTTATCGTTTGTCTACAGCTGCCGCTCAG 3096  
2937 CATGATGCCACCGTTCGATCAGCAAAATGGTTATCGCTTCGTGTACAGCTGCCGCTCTC 2996  
3097 CGCCGATCGGCTATTGATTTGAAGTACCCATTTACGTTAAACGACCGCGCTGGCGGAGAA 3156  
2997 GCGGACCAAGATTGTTAAATTTGAAGACACGCACTATATTGATATGCGACATTAGATCTCGA 3056  
3157 CACCGCTCTCAGACATCGCGACTATGCCAATCAGCAAGCTCGAGCTGAGTACGCT 3216  
3057 ATGCGCGCGCAAAATATTGCGGACTATGCCGCGCAACAGGGTTGGAGCTTCAGACACT 3116  
3217 GCTGCTGAAGACGACGGCATATTACCGATTACCCTGAGCGGCAACATCGATTCGATTCTG 3276  
3117 GCTGCGAAGAAACAGGGCGCTTACCATTACTCTGTGCGGCAATGCCAGCATTTCTG 3176  
3277 GCAACAGCAGCGCGGCGCAAGCTGAGCGGCTTGGCGCGCGGCTGTTTATGAGCACCAC 3336  
3177 G- ---CAGCAGCGCGCCCTGGGCTGTAGTGGATTACGTCGCGTCTGTTCCATCTCTACCAC 3233  
3337 CGGTTACTCTTTCGCTCGCGCTAGCGGCTAGCGAGTTGTTAGCAGCGCTGTTGCCAC 3396  
3234 CGGCTATTACCTGCGCTGGGCTTGGCTGGCGGACCGCTGAGTGCACTTGATGCTT 3293  
3397 CGATGCCCTCAGCTCAGCCAACTATCGAAACGCTTTGCGGTCAGCAGTGGCGGAAACA 3456  
3294 TAGCTCGGCTCAATTCACCATGCTTACCGATTTTGGCCCGAGCGCTGCGACGCA 3353  
3457 GCGATTTTTCGCTGCTTAACCGCATGCTGTTTGGCGCGTAAAGCGCTGTTTACGCGG 3516  
3354 GGGCTTTTTCGCGCATGCTGAATCGCATGCTGTTTTAGCGGAGCCGCGGATTCACGCTG 3413  
3517 GCGGTGATGCAAGTTTTACCGCTCGATGCGGGTTAATTAGCGCTTTTACGCGG 3576  
3414 GCGGTTATGAGGTTTATGTTTACCTTGAAGATTAAATGCCCCGTTTATGCGGG 3473  
3577 GCAACTCGCTGCGCGATAAACCGGATTTCTGTGCGGCAAGCGCGTGGCCATCGG 3636  
3474 AAACTCAGCTGACCGATCGCTACGTTATCTGAGCGGCAAGCGCTGTTCCGCTATT 3533  
3637 TGAAGCGTGGCGGCTGTTGAATTTCTGTCGAACCGAGGAAGAAATAAATAAGACGACT 3696  
3534 AGCAGCATTTCAAGCCCATTTATGACGACTCATPCTGTTAAAGAGCGACTCATGAAACCAACT 3593  
3697 TATGTTGCGCAGGCTTTGGCGGCTGGCGCTGGCGATTTCGCTGCAAGCGCGGCG 3756  
3594 ACGGTAATTGTTGAGGCTTCGGTGGCTGGCACTGGCAATTCTCTACAGCTGCGGG 3653  
3757 ATACCAACCACTTACTCAGCAGCGGCAAAAACCGGCGGACGCGCTATGTTGTTGAG 3816  
3654 ATCCCGCTTACTGCTTGAACACGTTGAATAAACCAGCGCTGCGGCTTATGTTCTACGAG 3713  
3817 GACAGTGGCTTATCTTCGATGCGGACCCAGCGTGTATCCAGATCCCGAGCCCATCGAA 3876  
3714 GATCAGGGTTTACTTTGATGCGAGGCGGACGCTTATCACCGATCCCGAGTSCCATGAA 3773  
3877 GAGTTGTTTACGCTGGCAGGAAATTCGCTCAGCGATTACGTCGAGCTGATGCCGCTAACG 3936  
3774 GAATGTTTGCATCTGGCAGGAAACAGTTAAAGATATGTCGAACCTGCTGCCGCTTACG 3833  
3937 CCCTTCTATCCGCTGCTGGAAGATGGCAAAACAGCTTGATTAAGCAATAAATAAGCCG 3996  
3834 CCGTTTACCGCTGTTGGGAGTCAGGGAAGGCTTTTAAATTAAGTAAACGATCAACGATCA 3893  
3997 CTGCTGAGCAGCAGATCGCCACGTTCAATCCGCAAGATGTAGAGGCTATCGTCAATTT 4056  
3894 CCGCTCGAAGCGCAGATTACGAGTTTAAATCCCGCGGATGTGGAAGGTTATCGTCAGTTT 3953





6105 CAACCGTTTGGAGTAGCCCTTGGTGGAAATATAGCGGAATGCCAACGTTGATGCACCAG 6164  
6277 GCATCGTCACCATGAAGTAGCGCGCGGTACGTCGTCATTCGGGACCAATCCACTG 6336  
6165 CCCGTCGTGCACATAAATAGAGTAATCCATACGCCGCTACACTGCGCCAATCCACTG 6224  
6337 CAGCGGCCACATGCGCTTGCACACCGACATCAATCAAGATCGCGCAGTACCGCAAAACAC 6396  
6225 GAGCGGCCACATCTCTGTACTGCCAGATAAATCAGCAGATCGATATGACGCAAAAC 6284  
6397 CACCGCATAAAGATCGTTGAGTCAAACTTACCGTGTGCGGTTTCATGTCGACAGATG 6456  
6285 CACCGCATAAAGATCGTTAACTTCAACCGCACCTTACCGCGTTTCATGATGTGAAGATG 6344  
6457 CAGCGCCATCCCAACCGTCGATGATGTTTATGCGACAGCGCGCTACGATTTCCAT 6516  
6345 CCATCCCAACCGCGCGTGCATGATGTTTGTGTCGCGCAGCAATCCTTCCAT 6404  
6517 CACCACCGCTTGCACCAAGATAGCAGCTTCCATACCGCAGCAGATGTTTCCGTCAT 6576  
6405 GCCATCACGGTAACGAAACGATCAGGCGCATTCBAATCCACACATAATTTCTCCGT 6464  
6577 TGTGGAAGGGAAGTA 6593  
6465 AGAGACGCTCGGCAGCA 6481

RESULT 4  
AY166713  
LOCUS  
DEFINITION  
Pantoea stewartii zeaxanthin-digluco-2002  
complete sequence.  
AY166713  
AY166713.1 GI:27228289  
Pantoea stewartii  
Pantoea stewartii  
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
Enterobacteriaceae; Pantoea.  
1 (bases 1 to 6586)  
deSouza,M.L., Kollmann,S.R. and Schroeder,W.A.  
Carotenoid Biosynthesis (WO 02/079395 A2)  
Unpublished  
2 (bases 1 to 6586)  
deSouza,M.L., Kollmann,S.R. and Schroeder,W.A.  
Direct Submission  
Submitted (18-OCT-2002) Biotechnology Development Center, Cargill  
Inc., P.O. Box 5702, Minneapolis, MN 55440, USA  
Location/Qualifiers  
1. .6586  
/organism="Pantoea stewartii"  
/mol\_type="genomic DNA"  
/strain="ATCC 8200"  
/db\_xref="ATCC:8200"  
/db\_xref="taxon:66269"  
169. .6399  
/genes="zeaxanthin-digluco-2002 biosynthetic operon"  
169. .1080  
/gene="crtE"  
169. .1080  
/gene="crtB"  
/gene="crtY"  
/gene="crtZ"  
/codon\_start=1  
/transl\_table=11  
/product="geranylgeranyl pyrophosphate synthase"  
/protein\_id="AA085596.1"  
/db\_xref="GI:27228290"

1090. .2385  
/gene="crtX"  
1090. .2385  
/gene="crtX"  
/codon\_start=1  
/transl\_table=11  
/product="zeaxanthin glucosyl transferase"  
/protein\_id="AA085597.1"  
/db\_xref="GI:27228291"  
/translation="MSHPAVIAPPFSSHVRALQAOELVARGHRVTFQOQDCKALV  
TGSIGQTVGLQTHPPGSLSHLHLAAHPLGSPMLRLINEMAPTSDMLCELPAAFH  
ALQEGVTDQMEPAGAVVAASGLPFSVACALPNNRPGPLAIVMEFFETSDAAR  
ERYTSEKIDYMRHRDRVIAHACRGLAPREKHLHCFSPQLAQISOLIBELDPK  
ALPCCFVAGPLRQPGTSGSTSYFSPDPKPRIPASLTGLQGHRYGLFRTIAKACEE  
VDAOLLAHCGGLSAGELARGDIOVDFADQSAALSOAQITTHGGNMTVLDAI  
ASRPLPLALPAFQPGVASRIVVHGIGKASRFTTSHALARQIRSLTLNTDYPQRM  
KLAARLAGSTPAADIVEQAMTCQPLVSGQYATAL"  
2369. .3517  
/gene="crtY"  
2369. .3517  
/gene="crtY"  
/codon\_start=1  
/transl\_table=11  
/product="lycopene beta-cyclase"  
/protein\_id="AA085598.1"  
/db\_xref="GI:27228292"  
/translation="MQPHYDLVAGLANGLIALRLOOQHPDMRILIEAGPEAGN  
HTWSPHEBDLTNQHRTAPLVVHWPYQVRRHRHNSGYCVTSRHFAGILRQ  
QFGHLMHTAVSAVHAESVOLADGRIIHASTDVIGRGYTPDSALRVGFOAFIGEMQ  
LSAHPGLSPIMDATVDQNGRYFVTLPLSATALLEDTHYIDKANLOERARONI  
RDYAAQWMPQTLREEGALPITLTGDNRFQWQOQPCAGSLNAGLHFHTTGYSLP  
LAVALADLSALDVTSSVHOTIAHFAQWQOQGFPMNLNRLFLAGPAESWRVM  
QRFYGLPEDLIARFYAGKLTVDRLILSGKPPVPVFAALQAIMTTHR"  
3529. .5007  
/gene="crtI"  
3529. .5007  
/gene="crtI"  
/codon\_start=1  
/transl\_table=11  
/product="phytoene desaturase"  
/protein\_id="AA085599.1"  
/db\_xref="GI:27228293"  
/translation="MKPTTVIAGFGGLALAIRLOAGIPVILLLEQRPKPGRAVYQ  
EQGTFDAGPTVITDPSAIEELFALAGKQLQVLELLPVTFFYRICWSGKRVFNDQ  
QAOLEAQIQPNRDRVAGYRAFLDYSRVFNKGLGTVFLPFLSKDMLRAAPOLAKL  
QAMRSVSKVAGYIEDEHLRQAFSLVGNPFPATSIYTLIHALERMGWFPFRG  
GTGALVNGIKLFODLGEVVLNARVSHMETVGDKIQAQVLEDGRFRFTCAVSNADV  
VHTVDDLISOHPAAKQAKLOSKEMSNLFLVYFGLNHHDDLAHHTVCFGRVREL  
IHEIFNHDGLAEDFSLIHAPCVTDPSLAPGCGSYVLAIPVPLGTANLDWAVSGPR  
LRDRIFDYLEQHYMPLGRSQLVTHRMFTTFDFRDELNAWQSGAFSVEFILTQSAWFRP  
HNRDKHDLNLYLVGAGTHPGAGIPGVIGSAKATAGLMLEDLI"  
5043. .5933  
/gene="crtB"  
5043. .5933  
/gene="crtB"  
/codon\_start=1  
/transl\_table=11  
/product="phytoene synthase"  
/protein\_id="AA085600.1"  
/db\_xref="GI:27228294"  
/translation="NAVSKSFATASTLFDKATRSVLMVYACWCHCDVDDOTLGF  
FADQSPQMPORLQOQLEMKTRQAYAGSQMHEPAPAFQEVMAHDIAPAYAFHLEG  
FAMDVRETRYLTLDLRYCYHVGAVGLMMAQMGVNDATLDRADIADGLAFLQITNI  
ARDIVDDAQVGRCYLPSWEEGLTKANYAAPENRQALSIRAGRLVREAEPPYVSSM  
AGLAQLPLRSAAWAIATAQVYRKIGVKVEQAGKQAWDRHQSTSTAEKLTLLTASGQA  
VTSRMKTYPPRAHLWQRP"  
complement(5872. .6399)  
/gene="crtZ"  
complement(5872. .6399)  
/gene="crtZ"  
/codon\_start=1  
/transl\_table=11  
/product="beta-carotene hydroxylase"

/protein\_id="AA085601.1"  
/db\_xref="GI:2728295"  
/translation="MLMIAWNLIVFVVVMEVVAALAHKVIHMGWNGWHLSHHEPR  
KFAFVNDLYAVVFAIVSIALIYFGSTGIWPLQWAGMTAYGLLYFVWHDGLVHQWR  
PPRYTPRGYLRKLYMAHRMHAVRKEGCVSFGFLYAPPLSKLOATLIRHARSGA  
ARDEQDGVDTSSSGK"

ORIGIN		
Query Match		37.4%; Score 2620; DB 1; Length 6586;
Best Local Similarity		64.7%; Pred. No. 0;
Matches 4016; Conservative		0; Mismatches 2170; Indels 24; Gaps 7;
QY	373	CAGCTGCAGCAATTTACAGCGCATCTTGAACGCTTACTGCTCCGCGGACAGCAAGC 432
DB	220	CAGTTGCTGGCTGATATGATAGCCGCTTGATCAGTTACTGCCGTTTCAGGGTACGGG 279
QY	433	GATCGCTGCGTCCGCGATGCGTCCGGAACGCTGGCGCAGGCAAAACGTTATTCGTCCT 492
DB	280	GATTGTGGGTGCCGATGCTGAAGCACGCTGGCACCCGGGCANAGTATTCGTCG 339
QY	493	TTATTACTGCTGCTGGCAGCGCGGATATGGGTGGAGCTGACGCAAAATGGCGTTCTC 552
DB	340	ATGCTGCTGTTATTAAACAGCGCGCATCTTGGCTGTGCGATCAGTCAGGGGGATTACTG 399
QY	553	GATCTGCTGTGCTGAGTGAATGTGACGCGGATCGCTGATTTCTGATGACATTCCTC 612
DB	400	GATTTAGCTGCGGCTTGAAATGTGATGTGCTGCTGCTGATTTCTGATGATATGCC 459
QY	613	TCGATGGATAACCGCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 672
DB	460	TCGATGGACGATGCGCAGATGCTGCGGGCGTCCACCATTTCAACGCACTACGGTGA 519
QY	673	AACTGGGATTTCTCGCGCCATTCGCGTCTTACCGCGCATTTTGAAGTGAATGCCATT 732
DB	520	CATGTGGCGATTCGCGCGCGTCTGCTTACTCAGCAAAAGCTTTGGGGTGAATGCCGAG 579
QY	733	GCACCGGTTTCCCTGCGCATATTAATCTGAAGGATGTCTGAACTCTCCGCTCGCTC 792
DB	580	GCTGAAGGTTCTGACGCGGATGCCAAAACCTCGCGCGGTGTGCGAGCTGTCCAACG 639
QY	793	GGCTGTCAGGGCTTAGTCAAGGCAATTCAGAGATCTGCAGCAGGACGCGACAGAGCGC 852
DB	640	GGCATGCGGGTCTGGTTCAAGGCGCATTTTAGGACCTCTCGAAGGCGATAAACCCGCG 699
QY	853	AGCCCGGAAGCATGCCATGACCAACGAACTGAAACACGAGCTGCTGTTTCGCGCCACG 912
DB	700	AGCGCGGATGCCATCTGCTAACCAATCACTGTTTAAACCCAGCACGCTGTTTTCGCGTCA 759
QY	913	CTGCAATGGCGCGCATTCGCTGACGCTTACCGCAGGTGCGGCAAGACTTAGCTTC 972
DB	760	ACGCAATGGCGTCCATTGCGGCCAACGCTCTCGAAGCGCTGAGAACCTGCATCGT 819
QY	973	TTCCGCGCAGGATTTGGGCGAGCGCTTCACTGCTCGACGCTCGCGAGGTTGCAAA 1032
DB	820	TTCTCGCTCGATCTCGGCGAGCGCTTTCAGTTGCTTGAGATCTTACCGATGGCATGACC 879
QY	1033	CACACCGGTAAAGATGTGACAGGATCAGGGCAAAATCCAGCTGGTACAGATGCTCGGT 1092
DB	880	GATACCGGCAAGACATCAATCAGGATGCAAGTAAATCAACGCTGTGCTCAATTTATTAGGC 939
QY	1093	GCTGACGGCGGAGAGTGGCTGCGGATCACTGCGCAGCGGAGATGACACACTTGGC 1152
DB	940	TCAGGCGGGTGAAGAACGCTTGGGACAGCATTTGCGCTTGGCGAGTGAACACCTTTCC 999
QY	1153	TGCGCTGCTATCGGGCATGCCACTCGCCAAA---TATATGACGCGCTGTTTAAATCAA 1209
DB	1000	GGGGATGCCAAAAGCGCATTTCCACACCCAACTTTTATTCAGGCGCTGGTTGACAAA 1059
QY	1210	CAGCTAGCGATATTCAACTAGCGCGGCTCAGCCGCTGGGCGCACTTTGCGGTGATCGCGC 1269
DB	1060	AAACTCGCTCGCTCAGTTAAGATGCTGCA-----TGAGCCATTTTGGGGTATCGCAC 1114
QY	1270	CGCGCTCTACAGCCACTTTTACGGGTTGACGGGCTTAGCACAACGCTGCTGGCGCGCG 1329

DB	1115	CGCCCTTTTTCAGCCATGTTTCGGCTCTGCAAAACCTTGTCTCAGGAATTAGTGGCCGCG 1174
QY	1330	GCCATCGCATCACATTCATCCAGCAAGCCGATGCCCGCACTTTGCTTAGCGACGAACGCA 1389
DB	1175	GTCATCGTTAGCTTTTTCAGCAACATGACTGCAAAAGCGCTGCTAAACGGCAGCGATA 1234
QY	1390	TCGATTTGTTGCGCTGCGCCCAACAGACGATCTGCGCGTTTCGTCGGGCGCGCTGTTCG 1449
DB	1235	TCGGATTCAGACCGTTCGACTGCAAAAGCATCTCCCGGTTTCCTTATTCGACCTGCTGC 1294
QY	1450	ATCGCTGCTGCTGCGCGCGCTGCTGCTGTTTCGCTGATCAGCATCTCGCGTCTCT 1509
DB	1295	ACCTGGCGCGCACCCACTCGAACCTCGATGTTAGACTGATCAATGAATGGCACGTA 1354
QY	1510	GCACGATATGCTGTGCGCGCAATGCTGCGGTACTGAAAGCATTTGAACATTCGATGGCG 1569
DB	1355	CCAGCATATGCTTTTTCGCGGAACTGCCCCGCGCTTTTCATGCTGCTGATAGAGGCG 1414
QY	1570	TGATCGCGCAGAAATGGAAGCGGCGGCGATTTGGTCTGTGAAGCGCTGCACTGCTGCT 1629
DB	1415	TGATCGTTGATCAAAATGGAAGCGGCGAGTGCAGTAGTCGCAAGAGCGTCAGGCTCGCT 1474
QY	1630	TTGTTTCGCTGCGCTGCGCTTTCGCGTCAATCGTGAAGCGGGATTCGCTTGGCGTGA 1689
DB	1475	TTGTTTCGCTGCGCTGCGCTTTCGCGCTCAACCGGAAACCGGTTTTCCTCTGGCGGTA 1534
QY	1690	TGCTCTTCGTTTTCAGAGGATGCAAAAGCGCTGAAACGTTTTCAGGCCACGACGATA 1749
DB	1535	TGCTCTTCGAGTACGGCACCGCATGCGCTCGGAAAGCTATACACACGACGAAGAAA 1594
QY	1750	TCTATGATCGCATCATGCTGCTGTCAGCGGACGCTGATCCTCAACACGCGGGCGTTTA 1809
DB	1595	TTTATGATCTGCTGATGCGAGCTCAGCATCGTGTGATCGCATCATGATCAGAAATGG 1654
QY	1810	ATTTGACGAGCGGCGGATTAATCAGTGCCTGCTGCGCGCTGCGACAAATCAGCCAGA 1869
DB	1655	GTTTAGCCCCGCGTGAATACTGCAATCATGTTTCTCCACTGGCAAAATCAGCCAGT 1714
QY	1870	TGCTGCGCGCTTTGATTTTCCAGTCAAGCAACTGCGCGCTGCTGCTATCAGCGCTGGGCG 1929
DB	1715	TGATCCCGCACTGATTTTCCCGCAAGCGCTGCGACAGCTGCTTTCATGCGGTTGAC 1774
QY	1930	CACCTCGCG---CGCGGTTTCTCTGCGCGCTCCATGCGCGCTGCGCAGCGCTGCTGTC 1986
DB	1775	CGTTACGCAACCCCGAGGACGCGCGGCTCATCAACTTCTTATTTTCGCTCCCGGACA 1834
QY	1987	AGCGGTGTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2046
DB	1835	AACCCCGTATTTTTCCTGCTGCGCACCTTGCAGGGACATCGTTTATGGCTGTTTCAAGA 1894
QY	2047	ATCTGGCGCAGCGCTGCGCGCATGCTGCGCTATGCTGCTGATGCCCAATTTGCGGGGAT 2106
DB	1895	CCATCGCAAGCTGCGAAGGTTGGATGCGCAGTTACTGTTGGCACTACTGCGCGGCC 1954
QY	2107	TAAACGCGCAACAGACGATCAGCTGAGTCTGCTGCGCGCGCTGCGTGGTGAAGATTGG 2166
DB	1955	TCTAGCCACGACGAGGAGTGAATGCGCGCGGCGGGAATTCAGGCTTGTGGATTTTG 2014
QY	2167	TCGATCAGCGCGGACCTTACAGCAGCGGAGCTGTTTATCACTCATGCGCGGTTAAACA 2226
DB	2015	CCGATCAATCCGACGACTTTTCAAGGCAAGTTGAAATCACAATGCTGGGATGAATA 2074
QY	2227	GCGGCTGGAAGCACTGGAATTCGCTGCGTACGCTGCTGCGCTGCGGATTTGCTTTGATC 2286
DB	2075	CGGTACTGGAGCTTATGCTTCCCGCACACGCTACTGCGCTGCGCTGCGCTGCGCTTATC 2134
QY	2287	AGCCCGGCTGCGCGCGCATTTGATGCGCATGAGTTGGTTCGCGCGCATCAACGCTTTA 2346
DB	2135	AACCTGGGCTGGCATCAAGATTTGTTTATCATGGCATCGGCAAGGCTGCTGCTGCTT 2194
QY	2347	GCGGTGTTTCACTCACTGAGGAGCATCTGCAACAGCTGCTGCAACGCTGCTTACGCG 2406

Db 2195 CTACAGCCATCGCTGGCGCGCAGATTGATCGTCTGCTGCTAACAACCCGATTACCGC 2254  
Qy 2407 TACGATGTCAGCGATTACGCGCAGCTGACGCGCGAGCGGTTGCGAGCGTGC CGCG 2466  
Db 2255 AGCGTATGACAAAAATTCAGGCGCATTCGGTCTGCGAGCGGCAACACAGCGCGCGCG 2314  
Qy 2467 ACATCGTCGAGCAGCGCTGTGCCAGCAGCAAGTGTGTGTCGCGAGGCGACCTGATCGG 2526  
Db 2315 ATATTGTTGAACAGCGCATGCG - GACCTGTACGCCAGTACTCAGTGGCGAGGATTATGCA 2373  
Qy 2527 CACGCAATACGATGTGATTTGGTGGTGTGCTGGAATGCGGAATGCTGATTCGCTGCG 2586  
Db 2374 ACCGCACTATGATCTCATTCGTGGTGGTGGCGGTCTGGCTAATGCGCTTATGCGCTCG 2433  
Qy 2587 TCTCGCTCAATTGTCAGCCACAACTGAATGCTGTTGCTGGAGCGGATGSCATCGCG 2646  
Db 2434 GCTTCAGCAACAGCATCCGATATCGGATCTTGCTTAATGAGCGGGTCTGAGCGGG 2493  
Qy 2647 AGGCAATCATACCTGGTTCATCATCAGCGATCTCAGCGCCGAACTTCGCTGGCT 2706  
Db 2494 AGGGAACCATACCTGGTCTTTCACGAAGAGGATTTAACGCTGAATCAGCATCGCTGGAT 2553  
Qy 2707 GCACCGCTGATACCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2766  
Db 2554 AGCGCGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2613  
Qy 2767 CAATCTGACGGGGATTATTTGCTCATCGCATCAGCGGATTTGCGCCGCCATCTTTACGC 2826  
Db 2614 CCATGTGAACAGTGGTACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2673  
Qy 2827 GCGGATGGGTGACGATCTGTGGAACAAACAGCGGTAACAAGGTAAACCAACGAGGT 2886  
Db 2674 ACAGTTTGGACAAATTTATGCTGCTATACCGCGTTCATGCTGAATCGGT 2733  
Qy 2887 GACGCTGGCGGATGCGGTGAACTTGTGCGCAAGTGGTGGTGGTGGTGGTGGTGGTGG 2946  
Db 2734 CAGTTAGCGGATGCGCGGATTTATTCATGCGCAAGTACAGTGGTGGTGGTGGTGGTGG 2793  
Qy 2947 GCGGACGCGCATCTGACGCTGGGTTATCAGGTGTTCTTGGAAGAGTGGCAGCTGGC 3006  
Db 2794 GCCTGATCTGCACACTACGCTAGGATTCAGGCAATTTATCGGTGAGGATGGCACTGAG 2853  
Qy 3007 GAGCGCGCAGCGCTGACGAGCGGATCCTGATGATGCCACCGTGCATCAGCAAGCGGG 3066  
Db 2854 CCGCGCGCATGGTTTATCGTCAACGATTTATCATGATGCGAGCGGTGCTGATCAGCAAAATG 2913  
Qy 3067 TTATGTTTGTCTACGCTGCGCTCAGCGCGGATCGGCTATGATGATGATGATGATGATG 3126  
Db 2914 CTACCGCTTGTGTTATACCTGCGCGCTTTCCGCAACCGCATCTGCTGATCGAAGACACA 2973  
Qy 3127 TTACGTTAAACAGCGCGCGCTGGCGGAGAACACCGCTGCTGACGACATCGCGCATATGC 3186  
Db 2974 CTACATTCAGAGGCTAATCTTCAGCGCGAAGCGGCGGTGAGAACATTCGGGATTATGC 3033  
Qy 3187 CAATCAGCAAGCTGGA CGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3246  
Db 3034 TCGCGGACAGGGTGGCGGTTACAGACGTTGCTGCGGGAAGAAAGAGGTGCTATGCCCAT 3093  
Qy 3247 TACCTGTAGCGGCAACATCGATGATTTGCGCAACAGCAGCGCGGCGCAAGCGTGCAGCG 3306  
Db 3094 TACGTTAAAGGGCGGATATCGTCACTTTTGGCAACAGCAAC - - - CGCAAGCGCTGTAGCG 3150  
Qy 3307 CCGCGCGCGGGCTGTTTCATGCAACACCGGTTACTCTTTCGCGTCCGCGCGCTGGCGCT 3366  
Db 3151 ATTACGCGCGGGCTGTTTCATCGCAACACCGGCTACTCCCTACGCTCGCGTGGCGCT 3210  
Qy 3367 AGCGAGTTGTGAGCAGGCTGTTGCCACCGATGCGCTCAGCTCAGCGCAACATATGCA 3426  
Db 3211 GCGCGATGCTCTCAGCGCGGCTGGATGTGTTTACCTCTTCTCTGTTCAACAGACGATTCG 3270  
Qy 3427 ACCTTTCCCGCTCAGCAGTGGCGGCAACGCGATTTTTCGCTCTGCTTAACCGCATGCT 3486  
Db 3271 TCACTTTCCCGCAACGTTGGCAGCAACAGGGGTTTTTTCGCGATGCTGAATCGCATGTT 3330

Qy 3487 GTTTTGGCGGTAAAGCGCAGCGCTGGCGGTGATGCAACGTTTTTTTACCGCTCGA 3546  
Db 3331 GTTTTACCGGACCGGCGGAGTCAAGCTGGCGTGTGATGAGGCTTTCTATGGCTTACC 3390  
Qy 3547 TGCGGGTTAATAGCGCTTTTACGCGGCAACTGCGCTGCGGATATAAAACGCGAT 3606  
Db 3391 CGAGGATTTGATTCGCGCTTTTATGCGGAAAACTCACCGTAGCCGATCGCTACGAT 3450  
Qy 3607 TCTGTGGGCAAGCGCGGTCGCCATCGGTGAAGCGCTGCGCGCTGTTGAATCTGT 3666  
Db 3451 TCTGAGCGCAAGCGCGCTTCCCGTTTCGCGCATTTGAGGCAATATATGAGACTCA 3510  
Qy 3667 CGAACAGGGAAGAAAAATGAAACGCACTTATGTGATTTGCGCAGGCTTTGGGCGCTG 3726  
Db 3511 TCGTTGAGAGCGACTACATGAACCACTACGTTAATTTGTTGGGGCTTTGGTGGCTG 3570  
Qy 3727 GCGTGGCGATTCCCTGCAAGCGCGGCAACAACCACTTACTCGACGAGCGAC 3786  
Db 3571 GCACTGGCAATTCGTTTACAGCGCGAGGTATTCCTGTTTGTCTGTGAGCAGCGAC 3630  
Qy 3787 AAA CGGCGGCAAGCGGCTATGCTTTGAGGACAGTGGCTTTA CTTTCGATGCGGACCC 3846  
Db 3631 AAGCGGGTGGCGGGCTTATGTTTATCAGGAGCAGGGCTTTACTTTGATGCGAGGCT 3690  
Qy 3847 ACCTGTGATCAACGATCCAGCGCCATCGAAGAGTTGTTACGCTGGCAGGAAAAATCGCTC 3906  
Db 3691 ACCGTTATCACCGATCCAGCGCGATTTGAAGACTGTTTGTCTGCGCGGTAAACAGCTT 3750  
Qy 3907 AGCGATTAAGTCGAGCTGATGCGGTAACCGCTTCTATGCGCTGTGCTGCGGAGATGCG 3966  
Db 3751 AAGGATTAAGTCGAGCTGTTGCGGTCAGCGGCTTTTATCGCTGTGCTGCGGAGTCCGGC 3810  
Qy 3967 AAA CAGCTTGAATACGACAAATATCAGCGCTGCTGAGCAGCAGATCGCACCTCAAT 4026  
Db 3811 AAGTCTTCAATTAAGATAACGACCGGCGAGTTAGAGCGCAGATCAGCAGTTAAT 3870  
Qy 4027 CCGCAAGATGTAGAAGGCTATCGTCAATTTCTTCCTTATTCACGTGAAGTATTTAGAGAG 4086  
Db 3871 CCGCGCGATGTTGGGGTATTCAGAGCGTTCCTTGACTATTCGCTGCGGTATTCATGAG 3930  
Qy 4087 GGTATCTGAAAATCGCGCAGGTCGCTTCTGAGGTGCGTGAATGCTGCGCGTGGCG 4146  
Db 3931 GGCTATCTGAAGTCGCGACTGCTGCTTTTATCGTTTCAAGACATGCTTCGCGCGCG 3990  
Qy 4147 CCGCAGTTGGGAGCTCTGCAAGCATGCGCAGCTCTACAGCATGCTGGCGAATTTATT 4206  
Db 3991 CCCAGTTGGCAAGCTGCGAGGCAATGCGCAGCGTTTACAGTAAAGTTGCGGCTACAT 4050  
Qy 4207 CAGCAGATCATCTGCGCTCAGCGCTTTTCTTCCATCTCATTTGCTGGGCGGTAATCCT 4266  
Db 4051 GAGGATGAGCATCTTCGCGAGCGCTTTCTTTTCACTGCTTTAGTGGGGGGAATCCG 4110  
Qy 4267 TTTGCAACGTCATCGATCTATACTTAAATTCATGCTGAGAGCGTGAATGGGGCGTGG 4326  
Db 4111 TTTGCAACCTCGTCCATTTATACGCTGATTCACGCTTAGAAACGGGAAATGGGGCTGG 4170  
Qy 4327 TTTCCGCGCGGCAACCGCGCGCTGTCAGGCGCATGCGCGACTGTTTCGAGGACTTG 4386  
Db 4171 TTTCCAGCGGTGGAACCGGTGCGCTGGTCAATGGCATGATCAAGCTGTTTCAGGATCTG 4230  
Qy 4387 GCGCGCGAGCTGTTTACTGAATGCGGAAGTGAAGCAGCTGGAAACAGCGGCAATTCGCAT 4446  
Db 4231 GCGCGCAAGTCTGTGCTTAACGCCCGGTGCTAGTCAATGGAACCGTTGGGGCAAGATT 4290  
Qy 4447 AGCGCGTTCATGTAGAGGCGGAGCGCTTCATGCGCGCGCTGTTGGCTCCAAATGCC 4506  
Db 4291 CAGCGCTGCAAGTTGGAAGACGCGCAGCGGTTTGAACCTGCGCGGTGGCGTCAAGCT 4350  
Qy 4507 GACGTGTGATCACTACGACAAATCTCTTCGCCACCATCCGCTGGCAATGGAACGCTGCG 4566  
Db 4351 GATGTTGATACATCTATCGGATCTGCTCTCAGCATCCGCGAGCGCTTAAGCAGCG 4410

QY 4567 ACATCGCTGAAGGTAAAGCGATGAGCACTCGCTGTTTGTACTCTATTGTCCTGAAT 4626  
DB 4411 AAAAAATCCAAATCCAAAGGTATGATGAATCACTCACTGTTTGTACTCTATTGTCCTCAAC 4470  
QY 4627 CAGCGCATGAACAGCTCGCGCACACACCGCTCTGTTTGGCCCGGTTATCGTCAGTTG 4686  
DB 4471 CATCATCAGATCACTCGCCCATCATACCGTCTGTTTGGCCACGCTACCGTGAATG 4530  
QY 4687 ATCGATGAGATTTTCAACAGCAGCCAGCTGGCAGACGATTTTTCATCTTTACCTGACCGG 4746  
DB 4531 ATTCAAGAAATTTTAAACCATGATGCTGGCTGAGGATTTTTCGCTTTATTATACGCA 4590  
QY 4747 CCTCGACAGGATCCGCTCGTCGACCCCGCTGGCGAGCTTTTATGTTGTAGG 4806  
DB 4591 CCTGTGTACGGATCCGCTCACTGGCAGCGGAAGGTGGCGAGCTATTATGCTGGCG 4650  
QY 4807 CCGGTGCGCATCTCGGCACCGCTGACATCGACTGGCAACAGGAAGACCGCTTGGC 4866  
DB 4651 CTGTGTCACACTTAGGCAAGGGAACCTCGACTGGCGGTAGAGGACCCGACTGGC 4710  
QY 4867 GATCGAATTTTGTCTATCTGAGCAGCACTACATCGCGGATTAAGTCAGCAATTAAGT 4926  
DB 4711 GATCGTATTTTGTACTTACCTTGAGCAACATTAATACCTGGCTTGGCAAGCCAGTTGGT 4770  
QY 4927 ACACACAGATTTTACGCGCTTTGATTTTTCGGACACGCTGCATGCCATCACGCTCG 4986  
DB 4771 ACGCACCGTATGTTTACGCGCTTTCGATTTTCGCGACGAGCTCAATGCCCTGGCAAGGTTG 4830  
QY 4987 GGTGTTTCGCTGAGCCGATTTTGAAGCAAGCGCTGCTCGCCCGCATCAACCGCAT 5046  
DB 4831 GCTTCTCGTTGAACTTATCTGACCCAGAGCGCTGTTTCGACCAATTAACCGCGAT 4890  
QY 5047 GCCGATATCAGCAATCTCTATCTGTTGGTGGCGTACGCAATCCAGGCGCGGCTGCC 5106  
DB 4891 AAGCATTGATAATCTTTATCTGTTGGCGCAGCACCCATCTGGCGCGGCTTCCC 4950  
QY 5107 GCGGTATCGTTGCGCAAGCGCACCGCAGGCTGATGCTGGAGATGCGCCGCAATGA 5166  
DB 4951 GCGCTAATCGGCTGCGCAAGCGCACGCGAGCTTAATGCTGGAGGACCTGATTTGACGA 5010  
QY 5167 ATCGACAGCTTTTACTTGAAGCAAGTAAACCAACCATGCGGTGGCTCGAAGATTTG 5226  
DB 5011 AT---ACGTCATTAATCATGATCGGTGCAAAACCATGCGGTTGGCTGCAAAAGCTTTG 5067  
QY 5227 CCACCGCCCGCAAGCTTTGATGCAACGACGCGCGCAGCAGCTGATGCTGATGCGT 5286  
DB 5068 CGACTGCATCGAGCTTTTCGACGCGCAAAACCGCTCGCAGCGTGTGATGCTTTAGCAT 5127  
QY 5287 GGTGTGCTCATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5346  
DB 5128 GGTGCGCGCACTGCGACGAGCTTATGACGATCAAAACACTGGGCTTTTCATGCGCCGAC 5187  
QY 5347 ATGCGCTGCAAGACGCGCAGGACGATGACGATCTGCAAAATGAAACCCGCGCGCT 5406  
DB 5188 CTTCTTCGAGATGCTGAGCAGCGCTTACGAGCTTGAATGAAACCGCTCAGGCT 5247  
QY 5407 ACAGCGCGCGCACATGATGAACCGCGCTTTAGGCGCTTTCAGGAAGTGGCGATCATTC 5466  
DB 5248 ACGCGGTTTCGCAATGACAGCGCGCTTTTTCGCGCTTTCAGGAGGTTCGCGATGGCG 5307  
QY 5467 ACCAGTGGCGCAACAACTGGGCTTTGATCATCTGGAAGGCTTCGATGATGATGATGATGAT 5526  
DB 5308 ATGATATGCTCCGCTTACGCGTTTCGACCATCTGGAAGGTTTTCGATGATGATGATGATGAT 5367  
QY 5527 ACGAATTAACGAGCTTCGATGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5586  
DB 5368 AAACGCGCTACCTGACACTGAGCAGTACGCTGCGTTTATGCTATCAGCTGCGCGGTGTTG 5427  
QY 5587 TCGGTTTGTATGTCGCGCTGTAATGGGCTGCGCAGCAAGCGGTTCGATGATGATGATGATGAT 5646  
DB 5428 TGGGCTGATGATGGCGCAATTAATGGGCTTTCGCGATTAACGCCAGCTCGATGCGCT 5487  
QY 5647 GCGATTTAGGACTGGCGTTCCAGCTCACTAACTTGGCGCGGACATTTGTAGAAGATGCGG 5706

DB 5488 GCGATCTCGGCTGCTTTCAGTTTGAACAATTCGCGGTGATATTGTGACGATGCTC 5547  
QY 5707 AAAATGCTCGCTGCTATCTGCGCAATCTCGCTCGATCAGCGGGATTAACGCGCGGATA 5766  
DB 5548 AGTGGGCGCTGTTATCTGCTGAAAGCTGCTGGAAGAGAAAGACTGACGAAAGCGA 5607  
QY 5767 CGCTGACTGACCGCAACATGTCGAGCGCTCGCTCACTGGCAGCGGCTTTAGTGGCGG 5826  
DB 5608 ATTATGCTGCGCCAGAAAACCGCAGGCTTAAAGCCGTATCGCGGGGACTGGTACGGG 5667  
QY 5827 AGGCGAAACCTTATATCACTCGCGCGATCCGCTTACCGGTTTACCGTGGCTCGG 5886  
DB 5668 AAGCGAAACCTTATATCACTCAATGCGCGCTGTCGCAATTAACCTTACGCTCGG 5727  
QY 5887 CGTGGCCATCGCTACCGCTCGCGCTTTTATCGCAAAATTTGGCGTCAAAAGTTTCAGCAG 5946  
DB 5728 CTTGGCCATCGCGACAGCAAGCAGGTGTACCGTAAATTTGGCGTGAAGTTGAACAG 5787  
QY 5947 CCGGTGTCAGCGCTCGGATTTCACGGACGCGACCAAGTAAAGTTGAAAACTGGCGCTG 6006  
DB 5788 CCGGTAAGCAGCGCTGGGATCATCGCCAGTCCACGCTCCACCGCGGAAATTTAACGCTTT 5847  
QY 6007 TGGTAAAGGGCAGGTTTGGCATCACTTCGCGTGTCTCGTCTGAAACCGCTGCGG 6066  
DB 5848 TGCTGACGGCATCCGCTCAGGAGTTACTTCCGATGAAGACGATTCACCCCGCTCGT 5907  
QY 6067 CTGCTCTGCGCAGCGCTCTGTTGATTTTACGTCGCTGAGCGCTGGCGAGCGTGGCTTG 6126  
DB 5908 CTCATCTGCGCAGCGCC-----GATCAGCGCATGCTTCTCTCAGCGTGGCTG 5961  
QY 6127 CAGCTTATTCAGCGGTGGCGGTAGAGAAACCAACGACGACGAGCTTTCAGCGCGCGG 6186  
DB 5962 AAGTTAGATAACCGTGGCGGTACAGAAAACCAAGGACGACGAGCGCTCTTTTCCCCT 6021  
QY 6187 CACCGCATGATGATGCGGTGGCGCATCTAAGCGCTTAAGATAGCTTTTCGCGCGGAT 6246  
DB 6022 TACAGCATGATGATGATGCGGTGGCGCATCTAAGCGCTTTCAGGTAGCGCTTTCGCGGAT 6081  
QY 6247 ATAGCGGAACCGCGCAGCGTTGATGACACGAGCGCATCTGTCACCATGAAGTAGAGCGCGC 6306  
DB 6082 GTAGCGGAACCGCGCAGCGCTGTTGATCAGTCCGTCGTGGACCATATAAATACAGTAAAC 6141  
QY 6307 GTACGCTGCTATTCGCGCAACATCCACTGACGCGGCGCACATGCTTGCACACGACATA 6366  
DB 6142 ATAAGCGCTCATGCTGCAACCAATCCACTGGAGCGCGCAGATTCCTGTACTGCGCAAGTA 6201  
QY 6367 AATCAGCAATTCGCGAGTACCGCAACACCAACCGCATAAAGATCGTTGAGCTCAAACTT 6426  
DB 6202 AATCAGGCAATTCGACCAATGCGCGAATACACGCGCATAGATCGTTAACTTCAATGCG 6261  
QY 6427 ACCGCTGTCGCTTTCATGTCGACAGATGCCAGCGCCCATCCCAACCGTGCATGATGA 6486  
DB 6262 GCTTTACGCGGTTTCATGATGTAAGATGCCAGCGCCCAACCCAGCGCTGCATGATGA 6321  
QY 6487 TTTATGCGACAGCGCGCTTACGATTTTCATCAACACCGGTTGCCAAGATAGCAC 6546  
DB 6322 TTTATGTCGAGTGCAGCAACCACTTCCATGCGCAGCAAGCGGTGACAAACACGATCAGGCG 6381  
QY 6547 GTTCCATTAACGAGCATTTGTCGTCATT 6576  
DB 6382 ATTCCAAATCCAAACATAAATCTCTCCAGT 6411

## RESULT 5

ERWCRTA

LOCUS

DEFINITION

ERWCRTA 6965 bp DNA linear BCT 11-APR-2001  
Pantoea agglomerans CrtE (crtE), CrtX (crtX), dycopene cyclase  
(crtY), phytoene dehydrogenase (crtI), and prephytoene  
pyrophosphate synthase (crtB) genes, complete cds.

ACCESSION

M90698

VERSION

M90698.1

GI:148393

KEYWORDS





QY 876 CAACGAACTGAAACCGAGCGTCTGTTTCGCGCCAGCGCTGCAAAATGGCGGCGAATTGCGCGC 935  
DB 1480 GAAACACATTTAAACACGACACCTGTTTTCGCGCTATATGCGAGATGGCGCTCAATTTGTGGC 1539  
QY 936 TGACGCTTCACCGCAGGTGCGCAAGACTTACGTTCTTCGCGCCAGGATTTGGCGCAGGC 995  
DB 1540 GAATGCCCTCAGGAAACGAGCGGATTAACCTACACCGTTTTCGCTTGATCTCGGCCAAGC 1599  
QY 996 GTTTCAACTGCTGAGCAGCTTCGCGACGTTGCGAACAACAACCGGTAAAGATGTGACCA 1055  
DB 1600 GTTTCAACTGCTGAGCAGCTTAACGATGCGCATGCTGACACGGGTAAAGACATCAATCA 1659  
QY 1056 GGATCAGGCAAAATCCACGCTGGTACAGATGCTCGGTGCTGACGCGGGAACGTTGCGCT 1115  
DB 1660 GGATGAAGGCAAAATCGACCTGGTCAATTTATTAGGCTCCAGACAGTTGAAGAGCGTCT 1719  
QY 1116 GCGCATCACCTGCGCAGCGCAGATGACACCTTGCCTGCGCTGCAATCGCGGATCGC 1175  
DB 1720 GCGGCACAGTCTGCACTCGCCAGTGAGCACCTCTCAGCGCGGTGTCAAAACGGTCATTC 1779  
QY 1176 CACTCGCCCAATATATGCAAGCGCTGTTTAAATCAACAGCTAGCGATTAATCAACTGAGCGCG 1235  
DB 1780 TACTCAACAATTTATCCAGCGCTGTTTGACAAAACCTCGCTGCGCTCAGTTAAGGATG 1839  
QY 1236 GCTCAGCGGTGGCCACTTTCGGGTGATCGCGCGCGCTCTACAGCCACTTTCAAGCG 1295  
DB 1840 CTGCA-----TGAGCCACTTCGCGGATCGCTCGCGCTTTTACAGCCATGCTCGTGCA 1894  
QY 1296 TTGCAAGCGTTAGCAAAACGCTGCTGGCGCGCGCATCGCATCACATTTCAATCCAGAA 1355  
DB 1895 CTGGAACCTCGCTCAGAGTTAGTGGCTCGCGTATCAGGTGACCTTTATTCAAGNA 1954  
QY 1356 GCGGATGCGCGCACTTTGCTTAGCGACGAACGCATCGATTTTGTTCGCGTGGCCCAACAG 1415  
DB 1955 CAGGACACTAAACCACTCATTCAGCATCCCACTCGGATTTTCGTAGGAGAGAA 2014  
QY 1416 AGCGATCTGCGGTTCGCTGGCGCGCTGTTGCTATCGCGTGGCTCGCGCGGCGCTG 1475  
DB 2015 AGCCATTCGCTGCTGCTCATTTATCGCGCTGCTGCACTTAGCGGCGCATCTCGGGCGG 2074  
QY 1476 TCGCTGTTTCGCGTATCAGCATCTCGCGTCTCGCACCGATATGCTGTCGCGCAACTG 1535  
DB 2075 TCAATGCTTCGCTCATCAATGAATGGCGCGCACGAGGATATGCTGTCCTGCGGCTC 2134  
QY 1536 CTGCGGTAATGAAGCAATGAACATCGATGCGGTGATCGCGACGAATAAGGAAGCGCG 1595  
DB 2135 CCAGAGCGCTTTAAACGATTTGCGATAGATGGTGTATCGTTGATCAGATGAGCGACGA 2194  
QY 1596 GCGGATGCTGCTGAGCGCTGCATCTGCGGTTGTTTCGGTGGCTCGCGCTGCGG 1655  
DB 2195 GCGCTCTGCTGCTGAAGCGCTGGGCGTGGCGTTTGTCTCGGTTGCTGTCACCTCGG 2254  
QY 1656 GTCAATCGTGAAGCGGGATTCGCTTGGGTGATGCTCGCTTTCGACAGGATGAC 1715  
DB 2255 CTTAATCGCAACCGGATTTGCGCTTGGCGGTGATGCGCTTTTGAATAGGCAACGCGAT 2314  
QY 1716 AAAGCGCTGAACGTTTTCAGGCCACGACGATATCTATGATCGCATCATGCGTGTGAC 1775  
DB 2315 GCGCAGGTGAGCGCTATGACGCCAGCGAAATAATTTATGATGCTGGTGTGCTCGGCAT 2374  
QY 1776 GCGGAGTGATCTCAACACGCGGGGCTTTAATTTGACGAGGCGGCGGATTAAT 1835  
DB 2375 GACCGGCTATTTGCCCATCACAGTACAGGATGGGCTTAGCACCCCGGGATAAATCCAT 2434  
QY 1836 CAGTGCCTGCGCGCTGCAAAATCAGCCAGATGGTGGCGCTTTCGATTTTCCAGT 1895  
DB 2435 CAGTGTGTTTTCGCAATGGCAAAATCAGTACGCTTATTCCAGAACTGATTTTCCCGT 2494  
QY 1896 CAGCAACTCGCGCTGCTATCAGCGCTGGGGCCACTCCGCGCGCGCTTCTCTGCG 1955  
DB 2495 AAAGCGCTACCGCTGTTTTCATGCTGTTGGCGCGCTCGGAGAAACGACAGCATCA 2554  
QY 1956 CCGTCCATGCGCGCTGCGCAGCGTGGCT---CAGCGGTGGTTTATGCGCTGCTGGGT 2012

DB 2555 ACCTTTGCGCGCTTTATTTTTCGCGCTGAAAGCCCGCAATTTTCGCTTTCGCTGGGC 2614  
QY 2013 ACGTTCGAAGGCAATCGCTTCGCGCTGTTTCTGCACTCTGGCGCAGCGCTGCCCGCAGCTG 2072  
DB 2615 ACGTTGCAAGGTCATCGCTACGGTCTGTTTAAAGCCATCGTGAAGCCTGTGAAGAGATT 2674  
QY 2073 CGGCTATCGCTGCTGATCGCCATTTGTGGGGATTAAGCCGGAACAGACGATCAGCTG 2132  
DB 2675 GATGTCAGCTTTCTGTTGGCGCACTGTGGCGCTGACTGCTTTTCAGGCCGAAGAGCTG 2734  
QY 2133 GAGCTCGCTGCGCGCGGTGCTGACGATTTTCGTCGATCAGCGCGCAGCCCTACAGCAC 2192  
DB 2735 GCGCAAGCGCTCAACACGAGTTTGTGGATTTTGCAGATCAGTCGCGCGCTGTCGCG 2794  
QY 2193 GCGCAGCTGTTTATCACTCATCGCGGTAAACAGCGCGCTGGAAAGCACTTGAATTCGCT 2252  
DB 2795 GCGCAGCTTGCCATTTACCCATGGCGGATGAATACGGTCTGATGCCATTAATATCGC 2854  
QY 2253 ACGCGATGCTGCGCTGCGGATTTGTGATGACCCGCGGTGCGCGCGGCGGCAATGAG 2312  
DB 2855 ACGCGCTTTTAGCCCTTTCGTTGGCATTTGATCAGCGCGCGCTGCGCTCAGCTATTGT 2914  
QY 2313 TGGCATGAGCTTGTGCGCGCGCATCAGCTTTAGCCGTTTATCATCACTGAGCAGCAT 2372  
DB 2915 TCTCATGGGATAGGAACAGTGGCTCCGTTTACACAGCCATGCGCTGCGCGCGCAC 2974  
QY 2373 CTGCAACAGCTGCTGACCGAGCATGTTTACGCGCTACGATGCTCAGCGATTCAGCGCGCAG 2432  
DB 2975 ATGCTGTCATGCTGACCAACAGTGCAGCTACCGCAACGATGACGAAATCCAGACCGCC 3034  
QY 2433 CTGCAAGCGCGCAGGCGGTTCGCGCGCGCATGCTGCGAGCAGCGCTGTCGCGCAG 2492  
DB 3035 CTTCTGTTTAGCGCGCGCAATAATGCGCGCTGCGATATCGTTGAGCAGCGCATGCG-CAC 3093  
QY 2493 CAGCAAGTCTGCTGCGCGGAGCGACCTGATGCGCGCACGCAATACCATGATTTGTTGCTG 2552  
DB 3094 CGCGCAGCGCTTTTGAAGGAGGCGGCTATGCGACCGCATTTGATCTGATTTTAGTGG 3153  
QY 2553 GTGCTGAGCTGCGCAATGCTGTTGATGCTGCTGCTGCTGCTCAATTTGAGCCACCACTGA 2612  
DB 3154 GCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3213  
QY 2613 AATGCTGCTGAGAGCGATGCGCATGCGCGAGGCAATATACCTGCTGCTGCTGCTGCTGCT 2672  
DB 3214 GTATTTGCTTATCGAAGCG 3273  
QY 2673 ACAGCGATCTCAGCGCGCAACAACTTTCGCTGCTGCGCAACCGCTGATTTACCGTGGTGT 2732  
DB 3274 ACGCTGATTTAACCGAGAGCCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3333  
QY 2733 CAGGTTATCAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2792  
DB 3334 CCGACTATCAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3393  
QY 2793 TCGCATCAGGCGATTTTCGCGCGCATTTTACGCGCGATGGGTGACGATCTGTGAGCAA 2852  
DB 3394 TGACTTCGAGCGTTTTCGCGAGGTTTACAAAGAGTTTGGTTCAGCACTTATGGAATTA 3453  
QY 2853 ACACAGCGTACACAGGTAAACCCAGCGAGTACGCTGCGCGATGCGCGCTGCGCTGCGCTGCTG 2912  
DB 3454 GCGCGCGGTGCGAGAGGTTCAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3513  
QY 2913 CTGCGCAAGTGGTGAATGATGCTGCGCGCTGCGCGCGCATGCGCGCGCATCTGCGAGCTGGGT 2972  
DB 3514 GTGCCAGCGCGGTGATGATGCGCGGATTAACGCCAAATTCAGCGCTGAGCGTGGAT 3573  
QY 2973 ATCAGGTGTTCTTGAACAAGATGGCAGCTGCGCGAGCGCGCGCGCTGCGCGCGCGCGCG 3032  
DB 3574 TCCAGCGCTTATCGGACAGGAAATGGCGACTCAGTAAACCGCACGCTTATCTTACCCCA 3633  
QY 3033 TCCTGATGATGCGCGCTGCGATCAGCAAGCGGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3092

Db 3634 TTATTATGATGCCAGCTGCTGATCAGCAAAATGGTTATCGCTTTGTTTACAGCCTGCCGC 3693  
Qy TCAGCGCCGATCGGCTATTGATTGAAGATACCCATTAGCTTAACAGCCCGCTGGCGG 3152  
Db 3694 TATCTGCGACAGAACTGTTAATCGAAGATACCCACTATATCGATAACGGGACACTGGAAC 3753  
Qy 3153 AGAACACCGCTCGTCAGCACATCGCGGCTATGCGCAATCAGCAAGGCTGGAGCTGAGTA 3212  
Db 3754 CTGAACGGCGCGACAAATAATTCGGATTATGCTGCCAGCAGGATTGGCAGCTTCAGA 3813  
Qy 3213 CGCTGCTGCTGAAGACGCGCATATTACCGATTACCGTAGCGGCAACATCGATCGAT 3272  
Db 3814 CTCCTGCTTGTGAGGAACAGGCGCATTCGCCAATAACCTTAACGGGCGATAGCACCGCT 3873  
Qy 3273 TCTGCGAACAGCAGCGCGCCCAAGCGTGCAGCGGCTGCGCGCGGCTGTTTCAATGCCA 3332  
Db 3874 TTTGCGACGAACAGC ---CTTTAGCTGCGAGTGGACTGGTGCCGACTGTTTACCCCA 3930  
Qy 3333 CCACCGGTTACTCTTGGCGTCCGCTGGCGCTAGCGAGTGTGTAGCAGCGCTGTTGC 3392  
Db 3931 CAACCGGTTATTCACTCCCGCTTGGCTGGCGGATCGCGCTGAGTGGCTTGATG 3990  
Qy 3393 CCACCGATGCCCTCAGCTCAGCCAAACATATCGAAGCTTTGCCCGTCAGCAGTGGCGG 3452  
Db 3991 TCTTTACATCGTCTCAATTATCATAGGCGATTACCACCTTTGCCACAGCGCTGGCAGC 4050  
Qy 3453 AACAGCGATTTTTCGCTCTGCTGATAAACCGCATGCTGTTTGGCGGTAAGCGCGAGCAGC 3512  
Db 4051 AGCAACGCTTTTTCGCGATGTTGATCGCATGCTGTTTGTAGCGGACCCGCGGATTAC 4110  
Qy 3513 GCTGCGGCTGATGCAAGCTTTTACCGBCTGCGCTAGCGGGTTAATTAGCGCTTTTACG 3572  
Db 4111 GCTGCGGCTGTTATGCAAGCTTCTATGTTTACCTGAAGATTAACTCCCGTTTTTATG 4170  
Qy 3573 CCGGSCAACTGCGCTGCGGATAAACCGCGATTCTGTGCGGCAAGCGCGCTGCGCCA 3632  
Db 4171 CCGGCAAACTCACTGTCAGCGCTACGCTACGTTTCTGAGCGGCAAGCGCGCTGTTCCGG 4230  
Qy 3633 TCGGTTGAAGCGCTGCGCGCTGTTGAATTCGTGCAACCGGGAAGAAATAAGAACG 3692  
Db 4231 TATTAGCGCATTTGAGGCAATTATGACGACTCATCGTTAAAGCGGACACATGAAACC 4290  
Qy 3693 CACTTATGATTTGCGCAGGCTTTTGGCGGCTGCGCTGCGGATTCGCTGCAAGCGGC 3752  
Db 4291 AACTACGGTAAATTGCTGACGCTTTTGGTGGCTGCGCATTAGCAATTGCTGTCAGCGGC 4350  
Qy 3753 GGGCATACCAACACCTTACTCGACGCGCGGACAAACCGGCGGACGCGCTATGCTT 3812  
Db 4351 GGGGATCCCTGCTTACTGCTTGAACGACGCGCAAAACCGGTGCGCGGCTTATGCTA 4410  
Qy 3813 TGAGGACAGTGGCTTTTACCTTCGATGCGGACCCACCGGTGATCACCGATCCGAGCGCAT 3872  
Db 4411 TGAAGATCAGGATTCATCTTTTGTATGACGCGCTACCGTTATACGATCCGAGCGCAT 4470  
Qy 3873 CGAAGATGTTTCA CGCTGGCAGGAAATTCGCTCAGCGATTAACGTCGAGCTGATCCGGT 3932  
Db 4471 TGAAGAACTCTTTTACCTTGGCGGGAACACAGTTAAAGATTACGTTGAATGCTCGCGT 4530  
Qy 3933 AACGCCCTTCTATCGCTGCTGCTGGGAGATGGCAACAGCTTGATTAGCAATAATCA 3992  
Db 4531 TCGCGGCTTTATCGCTGCTGTTGGGAGTCAGGAAGGTTTTCAAATTACGATAACGATCA 4590  
Qy 3993 GCGCTGCTGAGCAGCAGATCGCACCGTTCAATCCGAAGATGTAGAAGGCTATCGTCA 4052  
Db 4591 GGCAAGCTTTGAGCGGAGATTCAGCAGTTTAAATCCACGCGATGTTGAAGGCTATCGTCA 4650  
Qy 4053 ATTTCTGCTAATTCAGCTGAAGATTTTAGAGGCTTATCTGAATTCGCGACCGTGGCC 4112  
Db 4651 GTTCTGGAATTAATCAGCTGCGGCTGTTTAAAGAGGCTATCTGAAGCTTGGCACCGTGGC 4710  
Qy 4113 GTTCTGAGGTCGCTGATCATCTCGGCTGCGCGCGGCTTGGGACGCTGTCAGAGCATG 4172  
Db 4711 TTTCTTGTGCTTACGGATATGCTTCCGCGCGCGCCCAACTGGCGAAACTGCGAGCATG 4770

Qy 4173 GCGAGCGCTTACAGCATGCTGGCGAAATTTATTTCAGGACGATCATCTGCGTCAGCGCTT 4232  
Db 4771 GCGAACCGTTTACAGTAAAGTTGCGAGCTACATTCAGAGATGAACATCTTCTGTCAGCGCTT 4830  
Qy 4233 TTCTTCCACTCATTTGCTGGTGGCGGTAATCTTTTTCGCAACGTCATCGATCTATACCTT 4292  
Db 4831 TTCTTTTCACTCGCTGTTAGTGGCGGCAATCCCTTTCGCGACCTCTCTCAATTTATACGCT 4890  
Qy 4293 AATTTCATGCTGAGCGTGAATGGGCGTGTGTTTTCGCGCGGCGACCGGCGCGCT 4352  
Db 4891 AATACATGCACTCGAAACGCAATGGGCGCTGTGTTTTCACGCGGTGTATACCGGGGCATT 4950  
Qy 4353 GGTGAGGCGCATGGCGGCTGTTTCGAGGACTTGGGCGCGGAGCTGTTTACTGAAATGCCGA 4412  
Db 4951 AGTGAAGGATGATAAAGCTGTTTCAGGACTGGTGGCGGAGTGGTACTGMAATGCCAA 5010  
Qy 4413 AGTGAAGCGCTGGAACACAGCGGCAATCGCAATTAGCGCGCTTCAAGTTAGAGGGCGAG 4472  
Db 5011 GGTCAAGCCCATGGAACACAGCGGCGGATACCAATTGAAGCCGTGCAATTTAGAGGACGAGC 5070  
Qy 4473 ACGCTTTCATGCGCGCGCTGTGGCTCAATGCGGACGTCGTGTGTCATACCTACGACAACT 4532  
Db 5071 CAGATTCCGACCCGCGCTGTGCTTCCAAATGCGGATGTGTTTCACTATTCGGGACCT 5130  
Qy 4533 GCTTCGCCACCATCCGCTGGCAATGAACGTCGCAATCGCTGAAAGCGTAAAGCGCATGAG 4592  
Db 5131 GTTAAGTCAGCATCTTGC CGCGGTGAAGCACTCGAATAAACTGCAAACTAAGCGCATGAG 5190  
Qy 4593 CAATCGCTGTTGTGTA TCTATATTTTGGCTGCAATCAGCGGCAATGAAGCTGCGGACCA 4652  
Db 5191 CAATCACTGTTGTTCTCTATTTTGGTCTGAATCAATCAATCAATCAGCATGCTGTCACCA 5250  
Qy 4653 CACGCTGTTTTCGCGCGCTTATCGTGAAGTTCGATCGATGAGATTTTCAATCATGACGG 4712  
Db 5251 TACGTTGTTTTCGCGCGCTTACCTGAGTGTGATTCATGAAATTTTCAATCATGACGG 5310  
Qy 4713 GCTGCGACAGATTTTTCATCTTTCGCGCGCTTACCTGAGCGGCTTACGTCGCTGCGC 4772  
Db 5311 ACTGCGACAGATTTTCTACCTTTCGCGCGCTTACCTGAGCGGCTTGTGTCAGGATTCGTCATGCG 5370  
Qy 4773 ACCGCGCGCTGCGCGAGCTTTTATGTTTATGCTGAGCGCGGTCGCGATCTCGGACCGCTGA 4832  
Db 5371 ACCGGAAGTTTTCGCGAGTTACTATGTTGTTAGCTCCCGTGCACATTTAGGACAGCGAA 5430  
Qy 4833 CATGACCTGCGCAACAGGAAGCAGCGCTTCCGCGATCGAATTTTGTCTTATCTGAGCA 4892  
Db 5431 CCTGATGAGCGGTTGAAGGCGCGGACTCGCGGATCGTATTTTGTGATCTTGAACA 5490  
Qy 4893 GCATCTACATGCGCGGATTTACGTCAGCAATTAGTGACACACAGAAATGTTTACGCGCTTGA 4952  
Db 5491 ACATTTACATGCTGCTTTCGCTGAGCGAGTTGTTTACGCGGAAATGTTTACGCGCTTGA 5550  
Qy 4953 TTTTCGCGACAGCTGCTATGCGGCTTTCAGCGCTTTCGCGGTTTTCGCTGAGGACCGATTTTGAAC 5012  
Db 5551 TTTTCGCGACAGCTCAATGCTTATCAAGGCTCAGCTTTCGTTGTTGAGCGGCTTCTCAC 5610  
Qy 5013 GCAAGCGCTGCTTTCGCGCGGCTTACCGCGATGCGGATATCAGCAATCTCTATCTGCT 5072  
Db 5611 ACAGGCGCTTGTGTTTCGACCGCAACCGCGATAAACCATCAATTAATCTTTTACCTGCT 5670  
Qy 5073 GGGTCCCGGTACGATCCAGCGCGGCGTTCGCGGCGTGTATCGGTTTCGCGCAAGCGCAC 5132  
Db 5671 AGGTGCGGACCCATCCCGCGCGGCTTTCGCGGCTGATTCGCGCTCGGCTAAGCGCAAC 5730  
Qy 5133 CGCAGCGCTGATGCTGAGGATTCGCGCGGATGAATCGACAGCTTTTACTTGTAGGACGATA 5192  
Db 5731 TGCAGGTTAATGCTGGAGGA ---CTGATTGAATAATCCGCTGCTTACTCAATCATCGG 5787  
Qy 5193 ACGCAAAACATGCGGCTGCGGCTCGAAGAGTTTTCGCGCGCGCGCAAGCTGTTTGTATGCA 5252  
Db 5788 GTCGAAACGATGCTGTTGGCTTCGAAAGTTTTCGCGCTGCGCTCNAAGTTATTTGACGCA 5847



Qy	5253	CCGACGCGCGCAGCACGCTGATGCTGTATGCGTGGTGTGCTCAGTCGAGATGTGATTT	5311
Db	5848	AAAACCCGCGCAGTGTACTTATGCTACTCTCTGTGTGCGCCATTGTGACGACGTGATTT	5907
Qy	5313	GATGGGCAAACTGCTGGCGAAGCGGCACGCGCATGTCCTGCGAAGACGCGCAGGCACTG	5372
Db	5908	GAGCATCAAACCTCTGGGTGTTTCAAACCGATCAGCCGCTCATTTACAAACGCTTGACAAACGC	5967
Qy	5373	ATGCAGCATCTGCATAATTTGAAACCCGCGCGCCTTACACGCGCGCGCACATGAGTGAACCG	5432
Db	5968	CTGATGCAGTTGGAGATGAAACCGCGCCAGGCTATGTCGCGATGCGAGATGCGACAAACCG	6027
Qy	5433	CGCTTTAGGGCGTTTACAGGAAGTGGCGATCATTTACCAGCTGCCGCGCAACAACTGGCGTTT	5492
Db	6028	CGCTTTCTGCGTTTACAGGAAGTGGCGATGAGCCCATGATATTGCGCCGCTTTACGCGCTTT	6087
Qy	5493	GATCATCTGGAAGGCTTCGCTATGAGATGACGCAACGAACTTACGCGAGCTTCGATGAC	5552
Db	6088	GATCATCTGGAAGGCTTTGCCATGGATGTGCGTGAGCGGAGTACATTCAGTTGGACGAT	6147
Qy	5553	ACGCTGGTTACTGCTATCACTGCGCGGGCGTGTGCGTTTGTATGATGGCGCGCTAATG	5612
Db	6148	ACATTACGGTATTGCTATCATCTGTGACGGCTGTGCGCTTGTATGCGCGCAGATTATG	6207
Qy	5613	GGCGTGCGCAGGAAGCGGTCTCGATCAACGCTGCGCATTTAGAACTGGCGTTCAGCTC	5672
Db	6208	GGCGTACGCGATAACGCTACACTGGAACGCGCGCTGTGACCTTGGCGTGGCATTTCACTT	6267
Qy	5673	ACTTAACATTGCGCGGACATTTGAGAGATGCCGAAATGGTCGCTGCTATCTGCCGCA	5732
Db	6268	ACCAACATCCCGCGATATTGTGGAAGATGCGCATGTGGGTGCTGTTTATCTGCCGCGC	6327
Qy	5733	TCCTGGCTCGATCAGGCGGGATTACGCGCGATACGCTGACTGCACCGCAACATCGTGCA	5792
Db	6328	AGTTGGCTGGAAATGAAGNCTGAATTAAGAAATTTATGGGCACCGGAAACCGCCAG	6387
Qy	5793	CGCGTCGCCCTCACTGGCAGCGGTTTATGTGGCGGAGCGGAAACCTTATTACTACGCGCG	5852
Db	6388	CGCGTAGCGCGCATCGCCCGACGCTCTGGTGCAGGAAGCCGAAACCTTACTATTTTGTGCGCT	6447
Qy	5853	CGATCCGGTTTACCGGGTTTACCGCTCGGCTCGCGGTGGGCCATCGCTACGGCTCGCGGC	5912
Db	6448	ACGCGGGGACTTGGCCGGGCTGCCCTTCGCGCTCAGCCTGGGCCATTCGCAACAGGAAGCAG	6507
Qy	5913	GTTTATCGGAAATTTGGCGTCAAAGTTTCAGCACGCCCGGTGTGCACGCTTGCGGATTCACGG	5972
Db	6508	GTTTATCGTAATTTGGGTGNAAGTTTGAACGCGCAGCAAGAGCTTGGAACCATCGG	6567
Qy	5973	CAGCGCACCAAGTGAAGTGAACCTCGCGCTGCTGTGTGAAGGGGCAAGTTTGGCGATC	6032
Db	6568	CAGTCAACACACACACCCGAAATTTATCGCTTCTGTGTGATGGCATCCGCTCAGGCCATT	6627
Qy	6033	ACTTCGGTGTGCTCGTCTGTAACCGCGTCCGCTGGTCTGTGGCAGCGCTCTCTGTTGA	6092
Db	6628	ACTTCCCGGATGCGGCTCATTCGCGCTTCGCGCTGCCCATCTCTGGCAGCGCCGA-----	6682
Qy	6093	TTTTATCGCTGACGCTGGCGCAGCGTGGCTTCAGCTTATTCAGCGGTGGCGCGTAGA	6152
Db	6683	-TTTATCGGAATGTGCTTCTCTTAGCGTCCGCTGAAGTTTGTATAGCGCGGCGCATAGA	6741
Qy	6153	GGAAACCAAAACGACACGAGCTTTCAGCCCGCGCACCCGATGATGATGATGCGGTGGCCCA	6212
Db	6742	GMAAACCAAAAGAAACACAGCTTCTTTGCTCTTTACGGCATGATGATGATGCGGTGGGCCA	6801
Qy	6213	TGTATAGCGGTTAAGATAGGCTTTGGCGGGATATAGCGGAACCGGCGCAGGTTGATGCA	6272
Db	6802	TATACAAACCGCTTCAGGTAGCCTTTGGCGGGAATATCCGGCAACCGGCCAACGCTGATGCA	6861
Qy	6273	CCAGGCCATCGTGCAACCATGAAGTAGAGCGCGCGTACGTCGTCGTCATTTCGGCACCAATCC	6332
Db	6862	CCAGCCGCTGCTGAAACCATGAATATAGAGCAATTCATAGGCGGCTCATGCCCGCACCAATCC	6921
Qy	6333	ACTGCAGCGGCCACATGCTCTTGCAACCGGCACATAAATCAGCA	6374

Db 6922 ACTGAAGCGCCACATTCCGCTACTGCCACATAAATAGCA 6963

||||| ||||||||| || ||||||||| |||

RESULT 6  
 ERWCRTS  
 LOCUS  
 DEFINITION

12753 bp DNA linear BCT 11-APR-2001  
 Pantoea agglomerans pyrophosphate synthase (crtE), zeaxanthin  
 glucosyl transferase (crtX), and lycopene cyclase (crtY) genes,  
 complete cds; phytoene dehydrogenase (crtI) gene, partial cds;  
 phytoene synthase (crtB) and beta-carotene hydroxylase (crtZ)  
 genes, complete cds; and unknown genes.

ACCESSION  
 M87280 M99707  
 M87280.1 GI:148404

KEYWORDS  
 SOURCE  
 ORGANISM

Pantoea agglomerans  
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 Enterobacteriaceae; Pantoea.

REFERENCE  
 AUTHORS  
 TITLE

1 (bases 1 to 12753)  
 Hundle, B., Alberti, M., Nievelstein, V., Beyer, P., Kleinig, H.,  
 Armstrong, G.A., Burke, D.H. and Hearst, J.E.  
 Functional assignment of Erwinia herbicola Eho10 carotenoid genes  
 expressed in Escherichia coli

JOURNAL  
 MEDLINE  
 PUBMED

Mol. Gen. Genet. 245 (4), 406-416 (1994)  
 95107236  
 7808389

REFERENCE  
 AUTHORS  
 TITLE

2 (bases 1 to 12753)  
 Armstrong, G.A., Hundle, B.S. and Hearst, J.E.  
 Evolutionary conservation and structural similarities of carotenoid  
 biosynthesis gene products from photosynthetic and  
 nonphotosynthetic organisms

JOURNAL  
 MEDLINE  
 PUBMED

Meth. Enzymol. 214, 297-311 (1993)  
 93225848  
 8469144

REFERENCE  
 AUTHORS  
 TITLE

3 (bases 1 to 12753)  
 Hundle, B.S., O'Brien, D.A., Beyer, P., Kleinig, H. and Hearst, J.E.  
 In vitro expression and activity of lycopene cyclase and  
 beta-carotene hydroxylase from Erwinia herbicola

JOURNAL  
 MEDLINE  
 PUBMED

FEBS Lett. 315 (3), 329-334 (1993)  
 93138098  
 8422326

REFERENCE  
 AUTHORS  
 TITLE

4 (bases 1 to 12753)  
 Armstrong, G.A., Alberti, M. and Hearst, J.E.  
 Conserved enzymes mediate the early reactions of carotenoid  
 biosynthesis in nonphotosynthetic and photosynthetic prokaryotes

JOURNAL  
 MEDLINE  
 PUBMED

Proc. Natl. Acad. Sci. U.S.A. 87 (24), 9975-9979 (1990)  
 9108634  
 2263648

REFERENCE  
 AUTHORS  
 TITLE

5 (sites)  
 Hundle, B.S., O'Brien, D.A., Alberti, M., Beyer, P. and Hearst, J.E.  
 Functional expression of zeaxanthin glucosyltransferase from  
 Erwinia herbicola and a proposed uridine diphosphate binding site

JOURNAL  
 MEDLINE  
 PUBMED

Proc. Natl. Acad. Sci. U.S.A. 89 (19), 9321-9325 (1992)  
 93028456  
 1409639

FEATURES  
 source

Location/Qualifiers  
 1. .12753  
 /organism="Pantoea agglomerans"  
 /mol\_type="genomic DNA"  
 /strain="Eho10"  
 /db\_xref="taxon:549"  
 /tissue\_lib="ATCC 39368"  
 complement(1..1588)  
 /note="ORF2"  
 /citation=11  
 /codon\_start=1  
 /transl\_table=11  
 /product="unknown"  
 /protein\_id="AA064974.1"  
 /db\_xref="GI:551854"  
 /translation="MSELIHANAIEWRFRQILRLPEWGNLSLRVATCAPFTSAL  
 HALLTPESQAEILRDFAETILRNGLASVILNKLQALFYNQEGELLEWRQSRST

CDS

```
GIGASEKQDKYVSALKLDGREFKPLPGCKYQLTVRFEARQCEKLYGMGOYQQPWLDL
KGCSELAQRNSQAVPVPVQSGLGYGLLWNPALCEASFAXKNSQWTSRVTVQEMDIWI
TAGSDVMTROYAKATGTPPAAPFEIISGLWCKQRYRTQOEVLVAREVHERNPLS
VWVIDPFWPNQGTWCFDPVDPDPRAMVELESNGIRIMVSVWPTVARSPLPOMK
AKGLVTSDRGVQVNLDPNGNTTFPDATHPQARQVWDTVKKNYDFYKLGKFWLDBAE
PEYRADTQYKHAGPVLEVGNRIYPRDFAQGFYDGLRAEGENEIVNLVRCWSWAGSQR
YGVLAWSGDVSHFSFRNQLAAGLNMGLAGIPWTTDGGFQGNTHDPAPHELLIR
WFQWAVFCFPLRMHGYSRPRI"
complement (1596..2414)
/note="ORF3"
/citation=[1]
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAA64975.1"
/db_xref="GI:148406"
/translation="MLTWNNLLMVQLRQHRWLCFQVAYIFSSVGNGLTQTLILG
QLLRWHAPSTUTLITMLATLPGFISYILGAWMYFVYILGSAISYFMWMLAAHILGKA
AGSLAKLITRNFKVQIFGYCAVLGASIALFPAPKSVFLVELTPIITSLYQATT
TLMWMAADVADYGWSQGRMDGIIFSTFLAVLKLGMALGAIWGTILGFSGYVANA
PBOTSLMHCI VALFTVTVVPGLLSLAAFATLRWYKLDQTMQEBINLAKMQTIYKG"
3140..3451
/note="ORF4"
/citation=[1]
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAA64976.1"
/db_xref="GI:148407"
/translation="VALRSLSGTTGLHIPLPVFGNVFGWTVDEQSFSLDALLERGL
NAIDTADYSAWAPGNKGSEITILGKFWAHPGAREKITLFTKVGSDLKARQLMLRT
M"
3535..4458
/gene="crte"
3535..4458
/gene="crte"
/citation=[4]
/citation=[2]
/codon_start=1
/transl_table=11
/product="pyrophosphate synthase"
/protein_id="AAA64977.1"
/db_xref="GI:148408"
/translation="WVSGSKAGVSPHREIEVNRQSIDDHLAGLLPETSDQIVSLAMR
LRGMAPGRIRPMLLAARDLYQGSMTLIDLACAVELHTASLMLDMPQMDNAE
EKGQPTTHKFGSVAILASVGLSKAFGLIATAQDLPGERRAQAANLSTAVGVOG
LVLGFRDLNDALDRTDPAILSTWHLKTLGILFSAMLOIVATASASPSTRETLHAF
LDPCAGAILDLRDDHETGDKRNDKAGKSTLVNRLGADAAARQLRHHIDSADKHLT
FACPGGAIRQFMHLWFGHLLADWSFVWKIA"
4521..5564
/note="gene not found in Erwinia uredovora crt gene
cluster: ORF6"
/citation=[1]
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAA64978.1"
/db_xref="GI:148409"
/translation="MKDERLVQRKNDHLDIVLPPRAVTOASAGFERWRFTHCALPEL
NFSDTILFTFLNRQLQAPLLISSMTGGVRSRHNRLHAEAAQVILKIAMGVGSORVA
IESDAGLDKTLRLQAPDVPILLAGAAQLTKRGIDYARRAVEMISADALIVHNP
LQEAQLQGGDRDWRGLAAETLVRELVPVLVKEVGAGISRTVAGQLIDAGVTIDV
AGAGTQSWAAVGEERAETQQRVANVFADMGIPTAELVDIAEAWPQMPLIASGK
NGVDAAKALRLGACMVGQAAAVLSAGVSTEKVIDHFNVIIQLRVACFCFCTGSRSLSD
LKQADIRVYRDTp"
5561..6802
/gene="crtx"
5561..6802
/gene="crtx"
/citation=[3]
/citation=[1]
/codon_start=1
/transl_table=11
```

```
/product="zeaxanthin glucosyl transferase"
/protein_id="AAA64979.1"
/db_xref="GI:148410"
/translation="WSHEFAIVAPPLYSHAVAHALALEMAQGRHVRVTLTGNVASLAE
QETERVAFYPLPASVQOQRNVQOSNGNLRLLIAMSLSLTVLCQOLPAIILORLAVD
ALIVDEMPAGSLVAEALGLPFIACALPVRNRPGLPLPVMPHYAEDKALRRFQV
SHYIDALMYPPQPTILRHAQRFGLPERRRDECLPLAQISQSVFADDFPRRALPNC
FHYGALRYOPPPQVRSRSTPRIFASLGTLOGHRLRFQKIRACASVGAETVIAH
CDGLTPAQAOSLVAACGTVEVSFVDPQRYVAEANLVITHGGLNTVLDAALAAATPVYLA
PLSFDOPAAARLVNGLGRVSRFARQQTLDADLAEIAQLLGDETLHQRLATARQQLNDA
GGTPRAATLLEQAIGSESVS"
6799..7959
/gene="crty"
6799..7959
/gene="crty"
/citation=[3]
/citation=[2]
/citation=[1]
/codon_start=1
/transl_table=11
/product="lycopene cyclase"
/protein_id="AAA64980.1"
/db_xref="GI:148411"
/translation="MRDLILVGGGLANGLIARLRYRQYPQLNLLIIEAGEPQGNHTW
SFHEDDLTPGQHAWLAPLVAHWPGYEVQFPDLRRRLARGYISITSERFADLHQALG
ENIWLNGSVSEVLENSVRLANGEALLAGAVIDGRVTAASSAMQTYQLFLGQWRLTQ
PHGLTVPLMDATVAQQGVYFVYTLPLSADTLLIEDTRYANVPQRDDNALRQTVTDY
AHKSWOLAQLERETGCLPTLAGIDQALWADAPGVPRSGMRAGLPHPTTGYSLPLA
VALDADADSPLGSLVPLYOLTROFAERHWRQGFRLNMLFLAGREENRWRVMQR
FYGLPEPTVERFYAGRLSLEFDKARILITGKPPVPLGEAWRAALNHFPPDRDKG"
7956..99414
/gene="crti"
7956..99414
/gene="crti"
/note="putative; same bases, 11..1489, Genbank M38423"
/citation=[4]
/citation=[2]
/codon_start=1
/transl_table=11
/product="phytoene dehydrogenase"
/protein_id="AAA64981.1"
/db_xref="GI:551855"
/translation="MKKITVVIAGFGGLALIRLQAAGIPVLLIQRDKPGRAVYWH
DOGTDFAGPTVITDPTALEALFTLAGRMEYVRLLPVKFYLRCWESGKITLDYAND
SABEAQITQPNRDEVEGRFRPLAYSQAVFQEGYRLGSLVFLFRDMLRAGPQLKL
QAWQSVQSVSRFTEDHLRQAFSHLLVCGNFTTSSYITLHALEREWVWFPEG
CTGALNMGVXKLTDLGGEIELNARVEELVADNRVSOVLADGRI FOTDVAASNADV
VNTYKLLGHHPVGQKRAALERKSMNSLFVLYEGLNQPHSLAHHTICRGPYREL
IDELFTGSALADDFSLYHSPCVTDPSLAPGCSFYVLAAPVHLGNAPLWMAQEGPK
LRDRIFDYLEERYMPGLRSQVLTQRTIFTPADFDHDLDAHLGSAFSIEPLITQSAWFRP
HNRSDIANLYLVAGTHPGAGIPGVVASAKATASL"
9431..10360
/gene="crtb"
9431..10360
/gene="crtb"
/note="same bases, 1486..2415, Genbank M38423 but now
function has been reassigned"
/citation=[4]
/citation=[2]
/codon_start=1
```

Qy	1479	CTGTTTCGGTGCAGCATCTCGCTCTCGACCGATATGCTGCGCGCAACTGCT	1538
Db	5780	CTGCTGCGCGGTGATTGCGCGCATGTATCCCTCGCCAGATGCTCTGCCAGCGTGTCCC	5839
		Matches 3263; Conservative 0; Mismatches 1765; Indels 39; Gaps 4;	
Qy	1539	CGCGTACTGAAAGCATTTGAACATCGATGCGCGTGCACGAGAAATGGAAGCGCGGCG	1598
Db	5840	GCTATTCTACAGCGGCTGCGGTGCGACGCGTGTGTCGATGATGAGATGAGCGGCCGCGA	5899

Query Match 30.5%; Score 2135.2; DB 1; Length 12753;  
Best Local Similarity 64.4%; Pred. NO. 0;  
Matches 3263; Conservative 0; Mismatches 1765; Indels 39; Gaps 4;

QY 1599 GGAATTGGTCTGAAGCGCTGCATCTGCCGTTTGTTCGGTGGCTGGCCCTTGGCGGTC 1658  
DB 5900 AGCCTGGTTCGCCAGCGCTGGGACTACCAATTTATCTCTATTTGGCTGGCGCTGGCGGTC 5959  
QY 1659 AATCGTGAAGCCGGGATTCGGCTTCGGGTGATGCCCTTCCTCCGTTTGGCAAGGATGACAAA 1718  
DB 5960 AACCGGAGCCGGGTCTGCCGCTGCGGTGATGCCGTTTCACTACGCCGAGGATPAAGAGA 6019  
QY 1719 GCGCTGAAGCGTTTTCAGGCCAGCAGCGATATCTATGATFCGATCATGCGTTCGTCAACGCG 1778  
DB 6020 GCCCTCGCGCGTTTTCAGGTACAGCAACCGATCTAGGATCGCTGATGTACCCGACGCG 6079  
QY 1779 GAGGTGATCTCAAAACAGCGCGCGGTTTAAATTTGACGAGGCGGCGGATTAATCAG 1838  
DB 6080 CAGACGATCTCTGCCGACGCCAGCGCTTGTGTTGCGGAGCGAGCGCTCTCAGCAG 6139  
QY 1839 TGCCTGTGCGCGCTGGCAAAATCAGCAGATGTGTGCGCGCTTGTATTTTCCAGTCAG 1898  
DB 6140 TGTCTCTGCGCGCTGGCGGAGATTAGCAGTCCGTTCCGCTCTCGACTTCCACGCGCG 6199  
QY 1899 CAACTGCCCGCTGTCTATCAGCCGTTGGGCGCACTCCGCGCCCGGTTTCTCTCGCGCG 1958  
DB 6200 GCGCTGCCGAATCTTTTCACTACGTGGAGCACTGGCG-----TATCAGCGCGCG 6250  
QY 1959 TCCCATGCGCCCTGCGCAGCGCTGCTCAGCCGCTGTTATGCTCTGCTGGGTAGCTG 2018  
DB 6251 CCGCAGGTAGAACGTCTGCCACGCGCAGCGCGCGGATCTTTGCTCTGCTGGGCACTC 6310  
QY 2019 CAAAGCCATCGCTTCGCGCTGTTTCTGCACTGCGCGAGCGTGGCGCAGCTGCGGCTA 2078  
DB 6311 CAGGCGCACCGTCTAGCGCTGTTTCAAGATCGCCGCGCTGTGCCAGCGTGGGCGG 6370  
QY 2079 TCGTGGTGATCCGCCATTTGCGGGGATTAACGCCGAAACAGACGATCAGCTGGAGCTC 2138  
DB 6371 GAGGTGACCATTTGCCACTGCGATGSCCTGAGCCCGCCAGCGCTCGCTCTACGCC 6430  
QY 2139 GCTGCGCGCGTGGGTGAGGATTTCTGCACTGAGCGCGGAGCGCTCAGCAGCGCGAG 2198  
DB 6431 TCGCGCGGACGAGGTGTGTAGCTTTGTGCAACAGCGCGCTAGCTTGGCGAGGCTAAT 6490  
QY 2199 CTGTTTATCATTGATCGCGGTTTAAACAGCGCGCTGGAAGCACTGGAATGGGTACGCG 2258  
DB 6491 CTGGTGATCACCACGCGCGCTCAATACCGTACTGATGCTGCTGCTGCGCGGAGCGCG 6550  
QY 2259 ATGCTGGCGCTCCGATTTGATCAGCCCGCGGTGGCGCGCGATGAGTGGCAT 2318  
DB 6551 GTGCTGGCGGTGCCACTCTCTTTGCAACAGCCCGCGTGGCTGCGCGCTGTCTATAAC 6610  
QY 2319 GAGTTGGTGGCGCGCATCAGCTTTAGCGTGTCTCATCACTGAGCAGCATCTGCNA 2378  
DB 6611 GGGCTGGGTGCGCGGTTATCGCGCTTTGCCAGACAGCAGCGCTGGCGGATGAGATTGCC 6670  
QY 2379 CAGCTGTGACGACGATGTTACGCGCTACCGATGTGAGGATTTAGCGCGAGCTGCAG 2438  
DB 6671 CAACTGCTGGGGATGAGAGCTGCATCAGCGTCTGGGACGCGCGCGCAGCAGCTTAAC 6730  
QY 2439 CCGCAGCGGTTGTCAGCGTCCCGCGCATCGTCGAGCAGCGCGCTGTGCCAGCAGCAA 2498  
DB 6731 GAGCGCGGGGACGCCCGTGGCGGACCCCTGATTAACAGGCCAT----- 6777  
QY 2499 GTGCTGTCGGCGGAGCGACCTGATGCGCAGCGAATACGATGATTTGTTGGTGGTGTG 2558  
DB 6778 -----AGCAGGGAGTGAGAGCGTATCGTGAGGGATCTGATTTAGTTCGCGCGG 6826  
QY 2559 GACTGGCAATGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2618  
DB 6827 GCTTGGCCAAACGGGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6886  
QY 2619 TGTGCTGAGGAGGATGCGCATCCGGCAGGCAATCATACCTGCTGCTGCTGCTGCTGCTG 2678  
DB 6887 TGTGATCGAGCGCGGAGCAGCGCGCGGAAACCATACCTGCTGCTGCTGCTGCTGCTG 6946  
QY 2679 ATCTCAGCGCGGAAACAACTTCGCTGGCTGCTGCAACCGCTGATTAACCGTGGCTGCTG 2738

DB 6947 ATCTGACTCCCGGGCAGCAGCGCTTGGCTGGGCCCGCTGGTGCCCAACGCTGGCGGGCT 7006  
QY 2739 ATCAGGTGCGTTTCTCTGCGCTGGCCGCAATCTGGAGAGGGGATTAATTTTCATCGCAT 2798  
DB 7007 ATGAGTGCAGTTTCTCCGATCTTCCGCTGCGCTGCGCGCGGCTACTACTCAATACCT 7066  
QY 2799 CAGGCGATTTTGGCCCGCATCTTTACGCGGGGATGGGTGACGATCTGTGGAGAAACACAG 2858  
DB 7067 CAGAGCGCTTTTCCGAGGCCCTGCTCAGGCGCTGGGGAGAAATCTGGCTAAAATGTT 7126  
QY 2859 CCGTACAAACAGTAAACCCACGAGGTGACGCTGCGGATGCGCGTGAACCTTGTCTGCGC 2918  
DB 7127 CGGTGAGGAGGTGTTACCAATAGCTGGCTTGGCCAAACGTTGAGGCGCTGCTTGGCG 7186  
QY 2919 AAGTGGTGAATGATGCTGCGCGCTTGCAGCGACGCCACATCTGCAGCTGGGTATCAGG 2978  
DB 7187 GAGCGGTGATGACGAGCGCGGTGACCGCCAGTTCCGCGATGCAACCGGCTATCAGC 7246  
QY 2979 TGTTCCTTGGACAGAGTGGGAGCTGGGCGAGCGGCTGCGAGCAGCGGATCCTGA 3038  
DB 7247 TCTTTCTTGGTCAAGTGGCGGCTGACACAGCCCCAGCGCTGACCGTACCGATCCTGA 7306  
QY 3039 TGGATGCCACCGTGCATCAGCAAGCGGTTATGTTTGTCTACAGCTGCGCGCTCAGCG 3098  
DB 7307 TGGATGCCACCGTGGCGCAGCAGGAGTATCGCTTGTCTACAGCTGCGCGCTCTCCG 7366  
QY 3099 CCGATCGCGCTATTGATTAAGATACCCATTAAGTTAAACAGCGCCGCTGGCGGAGAA 3158  
DB 7367 CCGACACGCTGCTGATCAGGATAGCGCTAGCCCAATGTCCCGAGCGGTGATGATA 7426  
QY 3159 CCGCTGTCAGCACAATCCGCGACTATGCAATCAGCAAGGCTGGAGCTGAGTACGCTGC 3218  
DB 7427 CCCTACGCGCAGCGGTACCGACTATGCTCAGCAAAAGGCTGGAGCTGGCCAGCTTG 7486  
QY 3219 TGGTGAAGACAGCGCATATACCGATTACCTGAGCGGCAACATCGATCGATTCGCG 3278  
DB 7487 AACCGAGGAGACCGGCTGTCTGCGGATTAACCTGGCGGTGACATCAGGCTCTGTGG 7546  
QY 3279 AACAGCAGCGCGGCGCAGCGCTGCGCGCGCGGCTGTTTCAATGCAACACCGCG 3338  
DB 7547 CCGATGCGCGCGGCTGCGCGCTCGGGAATGCGGCTGGGCTATTTCAACCTACCATG 7606  
QY 3339 GTTACTCTTCGCTGCGCGCTGCGGCTAGCGGAGATTGGTAGCAGCGCTGTTGCCACCG 3398  
DB 7607 GCTATTCTGCTGCGCTGCGGCTGGGCTTGGCGACGCGATTGCCAGCAGCGCGCGCTG 7666  
QY 3399 ATGCCCTCACTCAGCACAATATCGAAACGCTTTGCCGCTCAGCAGTGGCGCGAAGCAG 3458  
DB 7667 GCAGGTTCCGCTCTATCAGCTCACCCGCGAGTTTCCGAAACGCCCATGGCGCAGCAG 7726  
QY 3459 GATTTTTCGCTGCTAAACCGCATGCTGTTTTCGCGCGGTAAGCGCGCAGCAGCGCTGC 3518  
DB 7727 GATTTTCGCTGCTGAAACCGGATGCTTTCCTGCGCGGCGCGAGGAGAACCGCTGC 7786  
QY 3519 GCGTATGAAACGTTTTCACCGGCTGATGCGCGGTTAATAGCCGCTTTTACGCGCGGCG 3578  
DB 7787 GCGTATGACGCGCTTTTATGGCTGCGGAGCCCAACCGTAGAGCGCTTTTACGCGCGCT 7846  
QY 3579 AACTCGCGCTGCGCATAAACCGCGATTCTGTGCGCAAGCGCGGCTGCGCATCGGTG 3638  
DB 7847 GCGCTCTCTTTGATTAAGCCCGCATTTTGAACGGGCAAGCCACCGGTTCCGCTGGCG 7906  
QY 3639 AAGCGCTGCGCGCTGTTGA-----ATTCTGTCGAACACAGGGAAGAAAAATGAACCGCAC 3695  
DB 7907 AAGCGTGGCGCGCGCTGACCATTTTCTGACAGACGAGATTAAGGATGAAGAAAC 7966  
QY 3696 TTATGTGATGCGCAGGCTTTGGCGCTTGGCGCTGCGCTGCGATTCGCTGCAAGCGCGCG 3755  
DB 7967 CGTTGTGATTTGGCGCAGGCTTTGGTGGCTTGGCGCTGCGCTGCGATTCGCTGACGCGCAG 8026  
QY 3756 CATACCAACCACTTACTCGAGCAGCGCAACCGGCGGACCGGCTGATGTGTTTGA 3815

Db 8027 GATCCCAACCGTACTGTGGAGCAGCGGCAAGCCCGCGGTGGGCTACGCTCGCA 8086  
Qy 3816 GGACAGTGGCTTTACCTTCGATGCGGACCCACCGTGTATCAACGATCCACGGCCATCGA 3875  
Db 8087 TGACCAGGGCTTTACCTTTGACGCGGGCCGACGCGTGTATCAACGATCCACGGCTTGA 8146  
Qy 3876 AGAGTTGTTACGCTGCGGAGGAAATCGCTCAGCGATTACGTCGAGCTGATGCGGGTAAC 3935  
Db 8147 GCGCTGTTTACCTTGGCGGCGAGCGCATGGAGATTACGTCAAGCTGTCTGCCGTAAA 8206  
Qy 3936 GCCCTTCTATCGCTGTCTGGAGATGGCAAAACAGCTTGATTAACGAATAATACGCC 3995  
Db 8207 ACCCTTCTACGACTCTGCTGGAGTCCGGAGAGAGCCCTCGACTATGTATTAACGACAGCGC 8266  
Qy 3996 GCTGCTGAGCAGATGCGCACGTTCAATCCCGCAAGATGATAGAGGCTATCGTCAATT 4055  
Db 8267 CGAGCTTGAGGCGCAGATTACCAAGTTCAACCCCGCGACGTCGAGGGCTACCGGCGCTT 8326  
Qy 4056 TCTTGCCTATTCACTGAAGTATTTAGAGAGGTTATCTGAAACTCGGCAAGGTCGCCGTT 4115  
Db 8327 TCTGGCTTACTCCAGGCGGTATTTCCAGAGGGATATTTGGCGCTCGGACAGCTGCCGTT 8386  
Qy 4116 TCTCAGGTGCTGATCATGCTGCGGTGCGCGCGAGTTGGAGCGTCTGCAAGCATGCGG 4175  
Db 8387 CCTCTCTTTCCGACATGCTCGCGCGCGGCGCGAGCTGTCTTAAGCTCCAGGCGTGCA 8446  
Qy 4176 CAGCGCTACAGCATGGTGGGAAATTTATTCAGGACGATCATCTGCGTCAGGCGTTTC 4235  
Db 8447 GAGCGCTTACCAAGTCCGCTTCGCGTATTTAGGATGAGCATCTGCGGAGGCTTCTC 8506  
Qy 4236 CTTCCACTCATTTGCTGGGCGGTAACTCTTTTGAACGTCATCGATCTATACCTTAAT 4295  
Db 8507 GTTCCACTCCCTGCTGGTAGGCGCAACCCCTTCAACCACTGCTCCAATCTACACCTGAT 8566  
Qy 4296 TCATGGCTGAGCGTGAATGGGCGGTGTGGTTTCGCGCGCGCGCACGCGCGCGCTGGT 4355  
Db 8567 CCACGCGCTTGAGCGGAGTGGGCGGTCTGGTTCCCTGAGGCGCGCACCGGCGCGCTGGT 8626  
Qy 4356 CAGGCGCATGGCGACTCTTCGAGACTTGGGCGCGAGCTGTACTGAAATGCCGAAGT 4415  
Db 8627 GAAAGGATGGTGAAGCTGTTTACCGATCTGGGCGGGGATCGAACTCAACGCCCGGT 8686  
Qy 4416 GAGCCAGCTGGAACACGCGCAATCGCAATTAGCGGCGGTTAGTTAGAGGCGGCGACG 4475  
Db 8687 CGAAGAGCTGGTGGCGCGATAACCGCGTAAGCCAGGTCGCGCTGGCGGATGGTCCGAT 8746  
Qy 4476 CTTTCGATGCGCGCTGTGGCTCCATGCGACGTTGGTGCATCTACGACAACTGCT 4535  
Db 8747 CTTTGCACCGACGCGCTGAGCTCGAAACGCTGACGTGGTGAACACCTATAAAGCTGCT 8806  
Qy 4536 TCGCCACCATCGCTGGCAATGAACGTCGACATCGCTGAAGCGTAAAGCGCATGAGCAA 4595  
Db 8807 CGGCCACCATTCGGTGGGCGAGAGCGGGCGGACGCGCTGGAGCGCAGAGCATGAGCAA 8866  
Qy 4596 CTCGCTGTTTGTACTATTTTGGCTGAATCAGCGCGATGAACAGCTCGCGCACACAC 4655  
Db 8867 CTCGCTGTTTGTCTACTTCCGCTGAAACAGCCCTCATTTCCAGCTGGCGCACCATAC 8926  
Qy 4656 CGTCTGTTTGGCGCGCTTATGTTGATGATGATGATGATGATGATGATGATGATGATGAT 4715  
Db 8927 CATCTGTTTGGTTCGCGCTACCGGAGCTGATCGACGAGATCTTTACCGGACGCGCT 8986  
Qy 4716 GGCAGACGATTTTCACTTTATCTGACGCGCTCGACGACGATTCGCTCGCTGCGACC 4775  
Db 8987 GCGCGATGACTTCTCGCTCTACCTGACCTCGCGCTCGCTGACGATCCCTCGCTCGCGC 9046  
Qy 4776 GCCCGCTGGGAGCTTTATGTTAGGCGCGGTGCGGCATCTCGGCACCGCTGACAT 4835  
Db 9047 TCCCGCTGCGCAGCTTCTACGTGCGCGCGGTGCGGCATCTTTGGCAACGCGCGCT 9106  
Qy 4836 CGACTGCAACAGGAAGGACCGCTTGGCGCATCGAAATTTTGTCTTATCTGAGCAGCA 4895  
Db 9107 GGACTGGCGCAGAGGGCGGAGCTGCGCGACCGCACTTTGACTACTCTTGAAGCG 9166

Qy 4896 CTATATGCGCGGATTACGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4955  
Db 9167 CTATATGCGCGGCTGGCTAGCAGCTGGTGAACCGAGCGATCTTTACCCCGCAGACTT 9226  
Qy 4956 TCAGCAGACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5015  
Db 9227 CCAGCAGACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9286  
Qy 5016 AAGGCGCTGCTCGCGCGCATTAACCGCATGCGGATGATGATGATGATGATGATGATGATGAT 5075  
Db 9287 AAGGCGCTGCTCGCGCGCATTAACCGCATGCGGATGATGATGATGATGATGATGATGATGAT 9346  
Qy 5076 TGCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5135  
Db 9347 GCGAGTACTCACTCGCGCGCATTAACCGCATGCGGATGATGATGATGATGATGATGATGAT 9406  
Qy 5136 CAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5195  
Db 9407 CAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9463  
Qy 5196 CAAACCATGCGGTGGCTCGAAGATTTCCGACCGCGCGCATGATGATGATGATGATGATGAT 5255  
Db 9464 CAGACCATGCGGTGGCTCGAAGATTTCCGACCGCGCGCATGATGATGATGATGATGATGAT 9523  
Qy 5256 ACAGCGCGCAGCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5315  
Db 9524 ACCGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9583  
Qy 5316 GGGCAACCGTGGCGGAGGCGCACGAGCATGATGATGATGATGATGATGATGATGATGAT 5375  
Db 9584 GACGAGACCCACGCGCTTTCGCGAGGAGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAG 9643  
Qy 5376 CAGCATCTGCAAAATTGAAACCCCGCGCGCTTACAGCGCGCGCACATGATGATGATGAT 5435  
Db 9644 GCCCGCTGCGCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9703  
Qy 5436 TTTAGGCGGTTTACGAGTGGCGCATCATTCACGAGCTGCGGCAACAACTGGCGGTTGAT 5495  
Db 9704 TTCGCTGCTTTACGAGGTTGGCGCTGATGATGATGATGATGATGATGATGATGATGAT 9763  
Qy 5496 CATCTGGAAGCTTTCGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5555  
Db 9764 CACTTCAGCGCTTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9823  
Qy 5556 CTGCGTTACTGCTATCAGCTCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5615  
Db 9824 CTGCGTACTGCTATCAGCTCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9883  
Qy 5616 GTGCGGAGCAGCGGTGCTGATCAGCTGCGGCTTATGAGCTGCGGCTGCTGCTGCTGCT 5675  
Db 9884 GTGCGGATGAGCGGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 9943  
Qy 5676 AACATTGCGCGCATTTGTAGAAAGTGCAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 5735  
Db 9944 AATATGCGCGGATATTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 10003  
Qy 5736 TGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5795  
Db 10004 TGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 10063  
Qy 5796 CTCGCTCACTGGGAGCGGTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 5855  
Db 10064 TGGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 10123  
Qy 5856 TCCGCTTACCGGTTTACCGCTGCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG 5915  
Db 10124 GCGCGCTACAGATCTGCGCGCGCTGCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG 10183  
Qy 5916 TATCGGCAAAATTGGCGTCAAAATTGATGATGATGATGATGATGATGATGATGATGAT 5975  
Db 10184 TACCGGAGATGCTGTTTAAAGTAAAGCGGCGGAGGAGCGGCTTGGATCGCGCGCAG 10243



Qy	2945	CAGCCGACGCCACATCTGACAGCTGGGTTATCAGTGTTCCTTCGACAGAGTGGCAGCTG	3000
Db	551	GCGCAGAAATTCAGCAGCTGAGCGTGGGCTTCGACGCGTTTATTTGGCAGGAATGGCCGATTG	610
Qy	3005	GGCGAGCGCAGCGGCTGACGACCGCATCTGATGATGCCACCGTCCGATCAGCAAGCG	3064
Db	611	AGCCACCCGATGGTTTATCGTCTCCATTTATCATGATGCGCAGGTTCGATCAGCAAAAT	670
Qy	3065	GGTTATCGTTTGTGTTACACGCTGCGCTCAGCGCCGATCGGCTATTGATGTAAGATACC	3124
Db	671	GGTTATCGCTTCGTGTACAGCTGCGCTCTCGCCGACCAAGATTGTGTTAAATTGAAGACAG	730
Qy	3125	CATTACGTTAAACAGCCGCGCTGGCGGGAACACCGCTCGTCAGCACATCGCCGACTAT	3184
Db	731	CACATATTGATATGCAATTAGATCTGTAATGCGCGCGGCAAAATTTTCGACTAT	790
Qy	3185	GCCAAATCAGCAAGGCTGGACGCTGAGTACGCTGCTGCTGAAGAGCAGCGCATTTACCG	3244
Db	791	GCFCGCAACAGGGTTGGCAGCTTCACACACTGCTGCAGAGAACAGGCGCGCTTACCC	850
Qy	3245	ATTACCTGAGCGGCAACATCGATCGATTTCTGGCAACAGCAGCGCGCCGAAGCGTGCAGC	3304
Db	851	ATTACTCTGTGGGCAATGCCAGCGCAATCTGG--CAGCAGCGCCCTTGGCCTGTAGT	907
Qy	3305	GGCCTGCGCGCGGCGCTTTCATGCCACACACCGGTTACTCTTCGCTCGCGCTGGCG	3364
Db	908	GGATTACGTGCGGCTGTGTTCCATCTTACCACCGGCTATTCACTGCGCTGGCGTTC	967
Qy	3365	CTAGCGAGTTGTGATGACAGCGCTGTGTCACACGATGCCCCCTCAGCTCAGCCAAACATTC	3424
Db	968	GTGCGCACGCGCTGAGTGCACTTGATGTCTTTACGTGCGGCTCAATTCACCATGCCAAT	1027
Qy	3425	GAACGCTTTGCGCTCAGCAGTGGCGGAACAGCGATTTTTTCGCTGCTGCTAAACCGCATG	3484
Db	1028	ACGCATTTTGGCCGCGAGCGCTGGCAGCAGCGGCTTTTTTCGCATGCTGATTCGCATG	1087
Qy	3485	CTGTTTTTGGCCGTAAGCCGACAGCAGCGCTGGCGCTGTGATGCAACGTTTTTACC	3544
Db	1088	CTGTTTTTACGCGAACCCGCGATTACGCTGCGCGGTTATGACAGCGTTTTTATGGTTTA	1147
Qy	3545	GATCCGGGTTAATTACGCGCTTTTACGCGGCACTGCGCTGCGCGATAAACCGCG	3604
Db	1148	CCTGAAGATTTAATTGCGCGTTTTTATGCGGGAACCTCAGCTGACCGATCGGCTACGT	1207
Qy	3605	ATTCTGTGCGCAACCGCGCTGCCCATCGGTGAAGCGCTGCGCGCTGTGTGAATCT	3664
Db	1208	ATTCTGAGCGCAAGCGCGCTGTTCCGCTATTAGCAGCATTTGCAAGCATTATGACGACT	1267
Qy	3665	GTCCAACGAGGAAGAAAATGAAACGCATTTATGATTTGGCGGAGGCTTTGGCGGCC	3724
Db	1268	CATCGTTAAAGAGCGACTACATGAACCAACTACGGTAATTGTTGTCAGGCTTCGGTGGCC	1327
Qy	3725	TGGCGCTGGCGAATTCGCTGCAAGCGCGGCGCATACCAACACCTTACTCGAGCAGCGCG	3784
Db	1328	TGGCACTGGCAATTCGTCTACAGCTGGGGATCCCGCTTACTGCTGTGAACACGCTG	1387
Qy	3785	ACAAACCGGGCGGACGCGCTTATGTGTTGAGAGCAGTGGCTTTACCTTCGATGCCGAGC	3844
Db	1388	ATAAACCGGGCGGTGGGCTTATGCTCAGAGGATCAGGGGTTTACCTTTGATGTCAGGCC	1447
Qy	3845	CCAGGTGATCACCGATCCAGCGCCATCGAAGATTGTTTCAGCTGGCAGGAAAATCGC	3904
Db	1448	CGACGGTTATCACCGATCCGAGTCCCATTTGAAGAACTGTTTTCCTGCGCAGGAACAGT	1507
Qy	3905	TCAGCGATTACGTGCGATGATGCCGCTTAAGCCCTTCTATCGCTGTGCTGGGAAGATG	3964
Db	1508	TAAAGAGTATGTCCAACTGTCGCGGTTACGCGGTTTTACCGGCTTGTGTGGAGTCAG	1567
Qy	3965	GCAACACGCTTGATTTAGACAAATATCAGCGCTGCTGGAGCAGCAGATCGCCACGTTCA	4024
Db	1568	GGAAAGGTTCTTTAAATTACGATAACGATCAAAACCCGGCTCGAAGCGCAATTCAGCAGTTTA	1627
Qy	4025	ATCCGCAAGATGTAGAAGGCTATCGTCAATTTCTTGGCTATTCTACGTGAAGTATTTAGAG	4084

[illegible]



Db 2708 CTGGCGTCAATCGGCTCGGCAAAAGCGACGACGAGTTTGTATGCTGGAGATC---TGATTT 2764  
Qy 5165 GAATCGACAGCCTTTACTTTGACAAAGTAACGCAAAACCATGGCGGTGGCTCGAAGAGTTT 5224  
Db 2765 GAATAATCGTCTTACTCAATCATGCGGTGAAACGATGGCAGTTGGCTCGAAGTTT 2824  
Qy 5225 CGCACCGCGCAAGCTGTTGATGACACGAGCGCGCGACGACGCTGATGCTGTATGC 5284  
Db 2825 TCGCAGAGCCTCAAAAGTTATTTGATGCAAAAACCGCGCGCAGCTACTGATGCTCTACGC 2884  
Qy 5285 GTGGTGTCTGCTCACTCGATGATGTGATGATGGGCAAAACGCTGGCGGAAGCGGCACGCA 5344  
Db 2885 CTGGTGGCGCATTTGTGACGATGTTATGACGATCAGACGCTGGGCTTTCAGCGCCGCGCA 2944  
Qy 5345 GCATCGCGTGAAGACGCGACGACGTAATGACGATCTGCAAAATGGAACCCGCGCGC 5404  
Db 2945 GCGTGCCTTACAAACCGCGCAACACGCTCTGATGCAACTTGAGATGAAAACGCGCCAGGC 3004  
Qy 5405 CTACAGCGCGCACATGATGAACCGCGGTTTAGGCGGTTTTCAGGAAGTGGCGATCAT 5464  
Db 3005 CTATCGAGATCGCAGATGCACGAACCGCGGTTTTCAGGAAGTGGCTATGCG 3064  
Qy 5465 TCACAGCTCGCGCAACAACTGGCGTTTGATCATCTGGAAGCTTTCGCTATGATGATGCAG 5524  
Db 3065 TCATGATATCGCCCGCTTACGCGTTTGATCATCTGGAAGCTTTCGCAATGATGATAG 3124  
Qy 5525 CAACGAACATTACCGGAGCTTCGATGACACGCTGCGTTTACTGCTATCACGTCGCGGCGT 5584  
Db 3125 CGAAGCGCAATACAGCAACTGATGATACGCTGCGCTATTCTCTATCATCGTTGACGGCT 3184  
Qy 5585 GTCGCTGTTGATGATGGCGGCTAATGGCGGTGCGGACGAGCGGTGCTGCTGATCACGC 5644  
Db 3185 TGTCGCTTGATGATGGCGCAATCATGGCGGTGCGGATACGCGCACGCTGACCGCGC 3244  
Qy 5645 CTGCGATTAGGACTGGCGTCCAGCTCACTAACTTGGCGCGCATTTGTAGAGATGC 5704  
Db 3245 CTGTGACCTTGGCTGGCATTTGATGACCAATATGCTCCGATATTTGGACGATGC 3304  
Qy 5705 CGAAAATGTCGCTGCTATCTGCCGCAATCTGGCTCGATCAGCGGGGATTAACGCGCGA 5764  
Db 3305 GCATCGGCGCGCTGTTATCTCCGCAAGCTGGCTGGAGCATGAAGGTCTGAACAAAGA 3364  
Qy 5765 TAGCTGACTGACCGCAACATCGTGCAGCGCTCGCTCACTGGCAGCGCGTTAGTGGC 5824  
Db 3365 GAATTATGCGGCACTGAAACCGCTGACGCGCTGAGCGGTATCGCCCGCTGTTGGTGA 3424  
Qy 5825 GGAGCGGAACCTATTATCACTCGCGCGATCCGTTTACCGGTTTACCGCTCGCTC 5884  
Db 3425 GGAACGAGAACCTTACTATTGTTGCTGCCACAGCGCGCTGGCAGGTTGCCCTCGCTC 3484  
Qy 5885 GCGTGGGCAATCGCTACGCTCGCGGCTTTTATCGGAAATGGCGTCAAAATTCAGCA 5944  
Db 3485 CGCCTGGGCAATCGCTACGCGCAAGCAGGTTTACCGGAAATAGGTGTCAAAATTTGAAC 3544  
Qy 5945 CGCCGCTGTGACCGCTGGATTCAGCGAGCGCACAGTAAGGTGAAAACCTGGCGT 6004  
Db 3545 GCGCGCTGACGAAGCTGGATTCAGCGGAGTCAACGACCGCGCGCAAAATTAACGCT 3604  
Qy 6005 GCTGTGTGAAAGGGCAGGTTTTCGCGATCACTTTCGCGTGTGCTCTGCTTCTGAAACCGCTCC 6064  
Db 3605 GCTGTGCGCGCTCTGCTGACGCGCTTACTTCCGATGCGGGCTCATCTCCCGCGCC 3664  
Qy 6065 GCGTGTCTGTGGCAGCGCTCTCTGTTGATTTTACGTGCTGACGCTGGCGCAGCGTGGT 6124  
Db 3665 TCGCATCTCTGCGAGCGCGC-----TCTAGCGCCATGCTTTCGCGAGCGCTGCC 3718  
Qy 6125 TCGACGTTATTCAGCGGTGGCGGTAGAGGAAACCAACGACACGAGCGCTTTCAGCGCG 6184  
Db 3719 TGAAGTTTTCAGAGGGCGCGCATAGAGGAAGCAAAAGAAACCAACCTTCTTTGGCC 3778  
Qy 6185 CGCACCGCATGATGATGCGGTGGCGCATGTATAGCGCTTAAGATAGCTTTTGGCGGG 6244  
Db 3779 CTGACGCGGTGATGATACGGTGGCGCATATACAAACCGCTTTCAGGTAGCCCTTGGCTGA 3838

Qy 6245 ATATAGCGGAACGCGCAGCGTTGATGACACGAGGCCCATCGTGACCAATGAAGTAGAGCGG 6304  
Db 3839 ATATAGCGGAATGGCCAAACGTTGATGACACGAGCGCTGTCACCAATAAATAGAGTAAT 3898  
Qy 6305 CCGTAGCTGCTGTTTCCGGCACCAATCCACTGCGAGCGGCCACATCGCTTGCACACCGCA 6364  
Db 3899 CCATAGCGGCTTACCTGCGCAATCCACTGAGCGGCCCAATCTCTGTACTGCCCGAGA 3958  
Qy 6365 TAAATCAGCAAAATCGCCAGTACCGCAAAACACCGCATAAAGATCGTTGAGCTCAAC 6424  
Db 3959 TAAATCAGCAGATCGATAATGCGCAAAACACCGCATAAAGATCGTTAATCTTCAAC 4018  
Qy 6425 TTACCGCTGCGGTTTATGTCGACAGATGCGAGCCCATCCCAACCGTGCATGATG 6484  
Db 4019 GCACCTTTACGCGGTTTATGATGTGAAAGATGCCATCCCAACCGCGCTGCTGATG 4078  
Qy 6485 TATTATGCGACGCGCGCTACGATTTCCATCACACACGCGTTGCCAAACAGATAAGC 6544  
Db 4079 TATTGTGTGCGAGTGCAGCAATCACTTCCATGCCAATCACGTAACGAAACAGATCAG 4138  
Qy 6545 ACGTTCCATAACCGAGCATTTGTCCTCATTTGTGAAAAGGGGAAGTA 6593  
Db 4139 GCAITCCAAATCCACAACATAATTTCTCGGTAGAGAGCTCTGGCAGCA 4187

RESULT 8  
BX571873 349107 bp DNA linear BCT 26-SEP-2003  
LOCUS Photorhabdus luminescens subsp. laumondii TT01 complete genome;  
segment 15/17.  
ACCESSION BX571873 BX470251  
VERSION BX571873.1 GI:36787441  
KEYWORDS complete genome.  
SOURCE Photorhabdus luminescens subsp. laumondii TT01  
ORGANISM Photorhabdus luminescens subsp. laumondii TT01  
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
Enterobacteriaceae; Photorhabdus.

REFERENCE  
1 Duchaud,E., Rusniok,C., Frangeul,L., Buchrieser,C., Taurit,S.,  
Bocs,S., Bouraux-Eude,C., Chandler,M., Dassa,E., Deroose,R.,  
Derzelle,S., Freyssinet,G., Gaudriault,S., Givaudan,A., Glaser,P.,  
Medigne,C., Lanols,A., Powell,K., Siguier,P., Wingate,V.,  
Zouine,M., Boemare,N., Danchin,A. and Kunst,F.  
Complete genome sequence of the entomopathogenic bacterium  
Photorhabdus luminescens  
Nat. Biotechnol. 11 (1) (2003) In press  
2 Duchaud,E., Frangeul,L., Rusniok,C. and Kunst,F.  
Direct Submission  
Submitted (23-APR-2003) L. Frangeul, Institut Pasteur, Genopole, 25  
rue du Docteur Roux, 75724 Paris Cedex 15, FRANCE. E-mail:  
lfrangeul@pasteur.fr, fkunst@pasteur.fr

FEATURES  
Location/Qualifiers  
1..349107  
/organism="Photorhabdus luminescens subsp. laumondii TT01"  
/mol\_type="genomic DNA"  
/strain="TT01"  
/db\_xref="taxon:243265"  
complement(80..3225)  
/gene="tccc1"  
/locus\_tag="plu4167"  
/function="unknown, probable insecticidal toxin"  
/note="Highly similar to insecticidal toxin complex  
protein Tccc of W14."  
/codon\_start=1  
/transl\_table=11  
/product="Insecticidal toxin complex protein Tccc1"  
/protein\_id="CAE15539.1"  
/db\_xref="GI:36787442"

```

/translaton="MSTPTDALTQTPTVSVLDNRGLSIRDGHRVWVIGGIDTRVT
RHOYDARGYLNSIDPRLYDAKQADNSVKPNFVWOYDIAGHALRTESVDAGRTTALND
TEGRVMTNNAITGVQTRREYEGNTLPGRLLSVSEQVPODETQVTERFIWAENTTAEK
EYNLSGLCIRHYDVTAGVRLMSQSLAGAILSQSHQWLAEGQEAWSGDETVWQMLA
SEVYTTQSTTNGALLTQDAKNIQRLAIDIAQOLKGSWITVKGREQVIVKSLSW
SAAHGKEEHNHGNGVITEYSYEPETQRLMGITTRAEESQSEARVLQDLRYEYDPVGN
VISHNDAEATFRNRQVBNPENRYVYDSLYQLMSATGREMANITQQSNQLPSPAMPV
PADNTSYNTQRTYSYDRGGNQLVRIHSGSPATQSYTTEMTVSSRNRNLTSLTITDP
TQDGLPDAGHGOKTLLI PQONLDMNRCGELORPVTPVRENSSDSEYRVSDGVRLLK
VSEQGTNGSTOVQRTVYLPGLRLRTGVADKTQNLQVITVGEAGRAOVRVHWESGK
PADIDNNVRYSYDNLSSQIELDSEQILSQERYYPYGGTAAWAAKQTEASYKFI
RYSGERNDATGLYGYRYYPQWGRWLSADPAGTVGLNRYVMRNNPIITLDHGL
APSPERNRNTWFASFLYRKPDEGMSMRGQKIGRAIAGGLAIAAGTAAIAGTAGA
ALPVLTVGAAYCAGTICAGLVNYSGLLEKGGALLARLVQGSKTLVQSAAGAAAGSA
AAYGARAQGVASAGAAVTCGASWINNADRIGCAIGAGSAGVCTIDTWLGTSTSLT
HEVAAAAGGAAGMTTGTQGSSTRAGIHAGITGYTGSWIGLGDVNASNAPGHLANVAG
YAAGLGAEMAINRVICGGFLGRLLGRAVSPYAGLARQLVHFSRARPVPEFIPIFVLGG
LAGGIGTGLUHRVMGDSWVSRGLSAGSGIDHLAGMIGNIGRARTVLTITTTGAINDYG
TSAIGAARVVISL"
complement (3220. .3225)
RBS
/gene="tccc1"
/locus_tag="plu4167"
/gene="tccc1"
complement (3345. .8051)
gene
/gene="tccc1"
/locus_tag="plu4168"
/gene="tccc1"
complement (3345. .8039)
CDS
/gene="tccc1"
/locus_tag="plu4168"
/function="Unknown, probable insecticidal toxin"
/Note="Highly similar to insecticidal toxin complex
protein TccB of W14 Putative transmembrane protein."
/codon_start=1
/transl_table=11
/product="Insecticidal toxin complex protein TccB1"
/protein_id="CAE16540.1"
/db_xref="GI:36787443"
/translaton="MLSTMEQLNESQSDALVTGYMNFVASTLKGVDGQPVTVDELVE
YLLIDPEVADEYETSRVQAIAISIQYMTVLVNGSEPGQRAEMLSTANERWDRNDSQYA
IWAAGAEVNADEYETSRVQAIAISIQYMTVLVNGSEPGQRAEMLSTANERWDRNDSQYA
VSNLVLGYINODKFDQAIYVIFGRTTKPVRYVYRWQMDLSKNRQDPAGNPVTPNCV
NDWQRTILPSGDTVLEHVPVNDRLVYAWVERDPAVOKDADGNIGKTHAYSIK
FGYKXDDITWAPNTTLLMTQADSESTQSRSLILDESKTLRKYNLLATDPSIDP
TEETSDNPGRLMGLVFRNFTGPHQNPVAVGYLYCDSAFNHHILRPLSKDFLPSF
YRDETNDQCLQFAYVDKXYVITKISGPTEDPENTGWGVDDLLKLTGTGAYYID
QDGLTLHTQTTNGFVNRHIFGYNDIVYDSKSGYFTWGSNEGFLYVHDGNYVTFH
NAAIYYPAGVGGSPVNGTWALQEVNAGWAI SPLLDHHTITVQGSYIAWEGTEPA
SYLPIPGTWMLDWSDKINFAISLNKLESVFTSPDWPTLTIKFNFSKIADNRKPYOE
INAETADGRNLFKHYSEQTFLASGATYPTTYTILSEADFPSTDPKNYLQVCLNVWDH
YDRSGRSKYSWKNFNVYVELQSKAPDAPRLIISKIDSKGLVQLVDFWTSSSLP
AKRLNTTTFVRLIEKANGLDLDYTLQADPSLEADLVTDGKSEPMDFGNSGLYF
WELFPHLPFLVATREFANQFSPAQKSLHYIPDPAMKKNPKNAPYMNVRPLVEVHSD
LSRHGLDPTDPTQAYAPHPVIYQKAVFIAYVSNLIAQGDWYRQLTRDGLTOARVYVN
LAAELIGPRPDVLSSTWTPQTLTAAQOKTALRDFEHLANSNTVLPALPGNNISH
LKLANGVFNEPLNVMLSHNDTLARLYNLHNLTVDGKPLSLPLYPAPADPVALLA
QQAQSGTLANGSGAILTPPYRFSAMLPRAYSAVGTLSFGNLLSLERSERISAGE
ELAQOQLDMSSYAILTQOQALDGLAARLALASQAQQRHDHYTLYQNNTSSAE
QLVMTQTSQAQSLISSSTGVOTASGALVFINI FGLADGSGRYEGVETAIAGLMAAG
QATSVVAERLATENRRRREWEQIOYQOAEVDALQOLDALAREKAATSQQA
KAQVQIRMTLTYLTRFTQATLYQLSGLSALYQAYDAVVSICLSAQACWOYEIG
DYATPIQGTWHDHVRGLQVETQLNLIOMEAYLYRHEERLNVITVYSKSLGD
DGFGLKTEGRVDFFLSEKLPFHYLQVIRKTVSVTLPIGLIPYQNVKATLTQTS
SSTLLTADINGVKRLDPTGEGDATHIVNLRASQOVALSSGINDAGSFELRLDER
YLSFEGTAVSKWTLTTPFRSVDHIDDKTLKADENQAALLANMDDVLVQVHYTACDGG
ASFANQVKKTLF"
complement (8046. .8051)
RBS
/gene="tccc1"
/locus_tag="plu4168"
/gene="tccc1"
complement (8134. .11034)
gene
/gene="tccc1"
/locus_tag="plu4169"
/gene="tccc1"
complement (8134. .11034)
CDS
/gene="tccc1"
/locus_tag="plu4169"

```

```

/function="Unknown, probable insecticidal toxin"
/Note="Highly similar to insecticidal toxin complex
protein TccA of W14."
/codon_start=1
/transl_table=11
/product="Insecticidal toxin complex protein TccA1"
/protein_id="CAE16541.1"
/db_xref="GI:36787444"
/translaton="MNOLASPLISRTBEIHNLPGLKTLDTGYTSFVDMVRPRFRFIRE
HRADLGSAREMYDLAVGHOVLHFRFRNSSEAVQFSLSRSPFSVSPDPVANOFPLDA
NTGKDXAPSGSPENDAPVAYLTHIQALDQKSGASATMNTLVERRPDI GALLIN
DAKNEVTPQQLVNEILSKAIKQJSLGKLEADNVAKLSTTRYPNNLPHYHQOIQIT
AQVLETTLDITLQPTLDLPQNFATAKGLDAGATAGALTRQLQMASQSPFQOQIM
VETVQGFYQLNYGDSLLTANSQDMTDLTSLTVPQVMEIMCSTVGGSTVVKSDN
VSSGDTATPAYGARFTHAGKPOAITLSRSGAEAFHATVNNLTDDKLDNRITVRL
QKWLNPYEDDLLVTSAMDACNTALSMNDNTLRMLGVKPKHQAKYVTKAPGAW
LRVTPAIPATPFPQOVNSVCTFETFPFVDNQDFVYTLTTCGDSGARVVKHISTALG
LNHQFLLADNATPQOQNTQSTLNCNLFVVSAPYRLANLARLTGTPNPFSCALVDR
LDAGTGVVQQQLAGKPIIMVQKDSPLAADLSLLQALSVIAQWQOQDLPFETILLL
LSSGSTFAQQTDDQNLNFIQVQWQLNSFTVSATLSRSGAPLVDTGHAIDWTFALLS
AGNSPLIDQVGLVTDAGTESVITMVVKTQSLSDDEKKAITLTNTLYAQOQTQGVV
VSLLAQTLNVQSPLPALLRWSGQTYQWLSATWALKOTVKIAADI PAGYLSLLREV
RBSLLTQOFTLSPAWQTLDPAYFGASAEVTDLSIMWLYTISRYSLDLQVKGAG
CTLEDDVIALYLRVANATPISQADAAQTLATLLGWETNELQAAHEVLGGIAKTTPLQD
TLRLQQAQSGTGLGVTQQQGYVLSRSDSTILWQSTGQAAVAGVSHVKGSH"
complement (11462. .11943)
gene
/locus_tag="plu4170"
/locus_tag="plu4170"
complement (11462. .11929)
CDS
/locus_tag="plu4170"
/function="Unknown"
/Note="unnamed protein and to Photorhabdus luminescens proteins"
immunity protein and to Photorhabdus luminescens proteins"
/codon_start=1
/transl_table=11
/product="CAE16542.1"
/protein_id="CAE16542.1"
/db_xref="GI:36787445"
/translaton="MAINIEALINSKTYQEI FDEGLIPYKTKPGYGPDSIDSLDM
AKEGIFLSENFENKVLTEITLFINHKKRPNFIPNKLPSLPLMSRKLHGHQFKPE
KSLPRKRLTKIEIGWTELYTLTLLDFRIPTSMQVDYDQERVLVTLFPTSEIRW"
complement (11938. .11943)
RBS
/locus_tag="plu4170"
/locus_tag="plu4170"
complement (12038. .12505)
gene
/locus_tag="plu4171"
/locus_tag="plu4171"
complement (12038. .12505)
CDS
/locus_tag="plu4171"
/function="Unknown"
/Note="unnamed protein product; Similar to pyocin S3
immunity protein and to Photorhabdus luminescens proteins"
immunity protein and to Photorhabdus luminescens proteins"
/codon_start=1
/transl_table=11
/product="CAE16543.1"
/protein_id="CAE16543.1"
/db_xref="GI:36787446"
/translaton="MTVNVNDALINSKGSYQEI FDEGLIPYKTKPGYGPDSIFALNM
AKEGMHLAFKEDKILFAVELLDOKEPYQFNELPSLPLKPLMTRVWVHEQFKPE
KALPRKFLKDXGTELYTLTLLDFRIPTSMQVDYDLEQVKSVAFLSTSEVLM"
complement (12597. .13081)
gene
/locus_tag="plu4172"
/locus_tag="plu4172"
complement (12597. .13064)
CDS
/locus_tag="plu4172"
/function="Unknown"
/Note="unnamed protein product; Similar to pyocin S3
immunity protein and to Photorhabdus luminescens proteins"
immunity protein and to Photorhabdus luminescens proteins"
/codon_start=1
/transl_table=11
/product="CAE16544.1"
/protein_id="CAE16544.1"
/db_xref="GI:36787447"
/translaton="MIDVIEIINSIGKTMADMLNKLIPKPKFPAFGYGRFSVEPE
TGLTLVFIISITVLSVQFTLINTVDGTYSQDMAPPLHAKMAIYRAFMGEFDS
SGPERPIGLVGGFDAYLRKLDQNTPEIRFLYQTLDRVALLFEQSVNND"
complement (13076. .13081)
RBS
/locus_tag="plu4172"
/locus_tag="plu4172"
complement (13174. .13652)
gene
/locus_tag="plu4173"

```



## CDS

complement(13174..13635) /locus_tag="plu4173" /function="Unknown" /note="unnamed protein product; Some similarities with pyocin S3 immunity protein and similar to Photorhabdus luminescens proteins" /codon_start=1 /transl_table=11 /protein_id="CAE16545.1"		Query Match Best Local Similarity 19.6%; Score 1374.8; DB 1; Length 349107; Matches 2211; Conservative 0; Mismatches 1347; Indels 8; Gaps 2;
QY	2532 AATACGATGTGATTTTGGTTCGGTCTGGACTGCGGAATGGCTTGATGCGTGCCTCTGC	2591
DB	202944 ACTGGGATCTGATTTCTGGTGGGAGGAGATTGGCCAAACGGACTGATCGCAATGCGTTCC	203003
QY	2592 GTCAATTGACGACCAACATGAATCGCTGTTCTGGAGAGCGATCGGCATCCGGCAGGCA	2651
DB	203004 AGCAGTGTAAACCAACATCTCGAGTGTGCTTATTGAAACACAGAAACAATAGGAGGCA	203063
QY	2652 ATCATACCTGTCGTTTCATCACAGGATCTCAGCGCGAACAACCTTCGCTGGCTGCAAC	2711
DB	203064 ATCACGTTGGTCATTTTCATCAACATGATCTTCTGAGCGGAAACATGATGGATAGCAC	203123
QY	2712 CGCTGATTACCGTGGTGGTTCAGGTATTCAGGTGGTTCCTGCGCTGCGCCGCAATC	2771
DB	203124 CGCTGATTACCTATCGTGGTCAGGTAGGACGTATTTTCCAGCATTTCAACGCACAT	203183
QY	2772 TGGACGGGATTAATTTGTCATCGCATCAGGCGATTTTGGCCGCGCATCTTTACGGCGGA	2831
DB	203184 TGCCCAATTCATATTTTCAGTATTCATCCAAACACTTTTGCAAGCATATCTCCATGCAATP	203243
QY	2832 TGGGTGACGATCTGTGGACAAACACAGCGGTACACAGTAAACCCAGCGAGTGACGC	2891
DB	203244 TGGGCGAAGTATACAGATCGTTTATGTCAGAGAGTGACTCCACAGAAAGTTTACT	203303
QY	2892 TGGCGGATGGCGGTGAATCTGTGCGCAAGTGGTGAATGATGGTGGCGGCTGCGACCGGA	2951
DB	203304 TACAGGACGGCTCGTCTCTAAGTGTGGCGAGTCATTTGATGGGAGAGGCTGGCGACCA	203363
QY	2952 CGCCACATCTGACGTGGGTTATCAGGTGTTCTTTGGACAGAGTGGCGAGCTGGCGCAGC	3011
DB	203364 GACCAATTTATAGGAGTGGCAACCCAGGCAATTTTGGTCAGCAATGGGAGCTGGAAGAT	203423
QY	3012 CGCAGGCTGACAGACGGATCTCTGTGATGATGCGCACCGTCGATCAGCAAGCGGTATC	3071
DB	203424 CGCATCTTTTAAACCCACCGATTTTAAATGGATACTAGTGTGGACAGGATACAGGTATTC	203483
QY	3072 GTTTTGTCTACAGCTGCGGCTCAGCGCGATCGGCTATTGATGAAAGATACCCATTACG	3131
DB	203484 GATTTATCTATGCTGCGGTTCTCTCAACTCGTCTGCTGATAGAGACACTCATTAGC	203543
QY	3132 TTAAACAGCCCGGCTGGCGGAGAAACACGCTGCTGATGACACATCGCCGACTATGCCAATC	3191
DB	203544 TTGATCGGGGCGCACCTGATAGGCTTTTGTGCGAGGCTACTATCGCAGACTACGCGCAAG	203603
QY	3192 AGCAAGGCTGGAGCTGATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3251
DB	203604 AACATGGATGGAACCTGGGTAACATCTTCGAGAGAGAGAGCGGTGTCTTCCAAATTCAC	203663
QY	3252 TGAGCGGCAACATCGATCGATTTCTGCGCAACAGCAGCGCGGCAAGGCTGACGCGGCTGC	3311
DB	203664 TTACGGGAGATTTTACCTCTTTCTGGCACAGCTAGCAGGACAGCCACCTGTGGGTTAC	203723
QY	3312 GCGCGGGCTGTTTCATGCAACACCGGTTATCTCTGCGCTGCGCGCTGCGCGCTGCGCG	3371
DB	203724 GTGACGCTTTGTTTACCCCAACAGAGGCTACTCTCTGCGCACACGCTATTTCCGTTGGCAG	203783
QY	3372 AGTTGGTAGCAGGCTGTTGGCCACCGATGCCCTCAGCGCTCAGCCACATATCGAACGCT	3431
DB	203784 ATCGTATTGTTGCTCTGCGGAGGCTTACCGATACCTCTTATTATTTACCTTACGGGATP	203843

QY	3432 TTGCCGCTCAGCAGTGGCGCGCAACAGCGAATTTTTCGGTCTGCTAAACCGCATGCTGTTTT	3491
DB	203844 ACGCAGCAACAGTGGCGCAACACAGCGCTTTTTCGGTCTTCTAAATCGCATGCTCTTCC	203903
QY	3492 TGGCGGTAAAGCCGACGAGCGCTGGCGGTGATGCAACGTTTTTACCGCTCGATGCGG	3551
DB	203904 TCGTGGGGATCCCAACACAGCTTGGCAGGTAATCAACGTTTCTATCAACTTTCCCAA	203963
QY	3552 GGTAAATAGCGCTTTTACGCGCGCAACTGCGCTCGCGGATAAAAACGCGGATTTCTGT	3611
DB	203964 ATCTGATTCAGCTTTTATGGGAGCAACTTAATTCGTCGACAGGCCCGGATTTCTCA	204023
QY	3612 GCGCAAGCCCGGTGCGCATCGGTGAAGCGCTGCGCGCTGTGA-----TCTGTG	3666
DB	204024 TAGGTAAACCAACCACTGCGCATAAAGGCTGCTTAAAGGCAATGTTAAACCAACAAGA	204083
QY	3667 CGAACCCGGAAGAAAAATGAACGCATTTATGATTGGCGGAGGCTTTGGCGGCTG	3726
DB	204084 AGCTTCAGGTTTTTATTTATTAAGCGCTGGTAAATTTGGTGTGGTGTG	204143
QY	3727 GCGCTGGCGATTGCGCTGCAAGCGCGGCGCATACCAACCACTTACTCGAGCAGCGCGAC	3786
DB	204144 GCATGGCAATAAGGCTCAGTCTGCGGGGATTCGACATGTTATTTTGGAGCAACGGAT	204203
QY	3787 AAACCGGCGGACGCGCTATGTGTTTCAGGACAGTGTGCTTACCTTCGATCCGAGCCC	3846
DB	204204 AAACCGGCGGACGCGCTATGTGTTTATAAGGAAACAGGGAATTCACCTTTGATGCGG	204263
QY	3847 ACGGTGATCAGGATCCAGCGCCATCGAAGTGTGTTTACGCTGCGGAGGAAATCGCTC	3906
DB	204264 ACCGTAATCAGCGCTCTTAATGTCAATGAAGAAATTTTATCCAGGCTGGTAAAGTATG	204323
QY	3907 AGCGATTACGTCGAGCTGATGCGGTAACGCCCTTCTATCGCTGTGCTGGGAGATGCG	3966
DB	204324 GCTGATACGTCGATTTACTTCCCGTACACCTTTTATCGCTCTGTTGGAGTCCGGT	204383
QY	3967 AAACAGCTTGATTACGCAATAATCAGCGCTGCTGGAGCAGCAGATCGCCACGTTCAAT	4026
DB	204384 AAGATGTTGATTTACGCAATGATCAGCAACATCTGGAAGCAACAATCCATACGTTCAAT	204443
QY	4027 CGCAGAGTGTAGAGGCTATCGTCAATTTCTTGCTATTACCTGAGAGTATTAGAGAG	4086
DB	204444 CCAAGAGATGTTAAGCGGTATCGACGCTTTCTAGACTATTCCGGAAGCTTTAATGAA	204503
QY	4087 GGTATCTGAAACTCGGACGCTGCGCTTTCTGACAGTGGCTGACATGCTGCGCTCGCG	4146
DB	204504 GGCTACCTGAAACTTGGAAACGGTGGCTTCTCTCTTTTCGCGACATGTTGAGTGGGCT	204563
QY	4147 CGCAGTTGGGACGTTCTGCAAGCATGGCGCAGCGCTTACAGCATGGTGGCGAAATTTAT	4206
DB	204564 CCACAGTTGTATAGCTCTACACGATGGCGTAGTGTATAGTCAGTGGCAGCGCTTATC	204623
QY	4207 CAGGACGATCATCTGCGTCAGCGGTTTTCTTCCACTCATTTCTGTTGGCGGTAATCT	4266
DB	204624 AAGGATGAGCGCTGCGCAGGCAATTTTCATTTCTCGCTATGATAGGAGTAACCCC	204693
QY	4267 TTTCGCAAGCTCATCGATCTATACCTTAAATCATGCGCTGGAGCGTGAATGGGCGGTGCG	4326
DB	204684 TTTCGTCATCTTCTATCTTACTTTGATCCACGCACTTGGAAACGAGATGGGAGTCTGG	204743
QY	4327 TTTCGCGCGCGGCGCACCGCGCGCTGGTGCAGGCGCATGGCGGATGTTTCGAGGACTTTG	4386
DB	204744 TTTCGCGGAGGAGAAACCAAGCGCTCTGGTTGAGGCAATGGTGAATTTGTTTACTGACAT	204803
QY	4387 GCGCGGAGCTTTTACTGATGCGCAAGTGAACGAGCTGGAACCAAGCGGCAATCGCAT	4446
DB	204804 GGTGAGAAATGAGCTTAATGCAAGGTAACCGCTTCCACCCATCGCAATCGGCT	204863
QY	4447 AGCGGCTTCAGTTAGAGGCGGACGACGCTTCGATGCGCGCTGTGGCTTCCCAATGCC	4506
DB	204864 ACAGGTGTCATTTAGTGTGAGGAGAAATATGACATGTGAGCTTGTGGCCTCAATGCC	204923

Qy	4507	GACGTGGTGCAATACCTACGACAAACTGCTTCGCCACCAATCCGCTGGCAATGAACGTGGC	4566
Db	204924	GATGTTATTTCATACCTATAAGACACCTGTTTAGGTACAGCATCCGGTCGGAATACCCCGCCT	204983
Qy	4567	ACATCGCTGAAGCGGTAGCGCATAGAGCAACTCGCTGTTGTGACTCTATTTTGGCTGAAT	4626
Db	204984	CGAACTCTGGTGGGTAAACGAATGAGCAACTCATCTGTCGTGCTTTATTTCTGGTCTGAAC	205043
Qy	4627	CAGCCGCATGAACAGCTCGCGCACCAACCGTCTGTTTGTGGCCGCGTTATCGTGAGTTG	4686
Db	205044	CATCACCATACGCAATTAAGCCATCATACAGTTTGTCTTGGTCACCGCTATAAGGAACTC	205103
Qy	4687	ATCGATGAGATTTTCAACAGCAGCCAGCTGGCGAGACGATTTTCACTTTACTCTGACACGG	4746
Db	205104	ATCGAAGATATTTTTTTATCATGACCGCATGTGAGAGGATTTTTCACTCTATCTCCATGCC	205163
Qy	4747	CCCTGACAGACGGATCCCTCGCTGCGCACCGCCGCTCGGCAGCGTTTTATGTGTGATGCG	4806
Db	205164	CCCTCTGTTACTGACCCATCCCTTGTCTCCCGAGGATGCGAAAGTTACTAGTTCGTGGCC	205223
Qy	4807	CCGGTCCGGATCTTCGGCACCCGCTGACATCGACTGGCAACAGGAAGAGACCGGCTTGGCC	4866
Db	205224	CCTGTACCACATTTTGGGACAGCTAAATTTAACTGGGATATAGAGGGGCGACGCTTACGT	205283
Qy	4867	GATCGAATTTTTTGCTTATCTCGAGCAGCACTACATGCCGGGATTAACGTCAGCAATTAGTG	4926
Db	205284	GACCGTATTTTCGGGTATCTGGAAAGTACTATATGCCCGGCTCTGCTCAAGCAGTTAGTC	205343
Qy	4927	ACACAGAAATGTTTATCGCGTTTGAATTTTCGGACACGCTGCAATGCCCATCAACGGCTCG	4986
Db	205344	GTTTCATCGTATTTTTTACCCCATTTGATTTTCTGTGACCAACTCAATGCCCATCTCGGTTTCG	205403
Qy	4987	CGCTTTTCGCTGGAGCCGATTTTGAACGAAAGCGCTCTGTTCCGCCCGCATAAACCGCAT	5046
Db	205404	GCTTCTCCTTTGAACCGTTGCTGACAAAGCGCTTTGGTTCGGACCAACAATCGGAC	205463
Qy	5047	GCCGATATCAGCAATCTCTATCTGTTGGTGGCGGTACGCAATCCAGGCGCGGCGTGGCC	5106
Db	205464	AACCGCATTTGATAATCTGTATCTCTGCGCGCAGGACGCACTCTCGAGCAGGTATTCCT	205523
Qy	5107	GGCGTGATCGTTTCGGCCAAAGCCACCGCAGGCTGCTGCTGGAGGATCGCGCCGAATGA	5166
Db	205524	GGCGTGATTTGATACGAAAAAGCCACCGCCACATTTGATGTTAGAGGATATCGCTAAATGA	205583
Qy	5167	ATCGACAGCCTTTACTTGAGCAAGTAAGCAAAACCATGGCGGTGGGCTCGAAGATTTTCG	5226
Db	205584	ATCCTGGC--TACTTAAACAGTTACTCAGATAATGGAGCAAGGCTCAAAAAGTTTTTG	205640
Qy	5227	CCACCGCCCAAGCTGTTTGAATGACCGACCGCGCGAGCAGCTGTGATGCTGTATGCGT	5286
Db	205641	CCAGTGTCAACCGACTTTTCGATACAGCAACACGACACAGCAATATGATGCTGTACGCT	205700
Qy	5287	GGTGTGCTCACTGCGATGATGTATTGATGGGCAAAACGCTGGCGGAAGCGCGCAACGACG	5346
Db	205701	GGTGGCGTTATTTGTGATGATATAATAGACGGCAGAGCTGGGGAGACAAATAAGCAGTG	205760
Qy	5347	ATGCGGTGGAAGACGGCAGGACGATATGACGATCTGCAAAATGAAACCCCGCGCGCT	5406
Db	205761	TTGATAAGATTATAGCGCCGCGAGAACTTCAAATGCTGCAATACCTGCAAAAAGCAGGCTT	205820
Qy	5407	ACAGCGGCGCGACATGGATGAACCGGCTTTTAGGGCGTTTTCAGGAAGTGGCGATCATTC	5466
Db	205821	ACGACGGTTTGGGATGACCGAAACCTGCTTTTGGAGCCTTTTACAGACGTAGCTCTGAGTA	205880
Qy	5467	ACCAGCTCGCAACAACTGGCGTTTGAATCATCTGGAAGGCTTCGCTATGGAATGACACGCA	5526
Db	205881	ATGAAATCCCTCAGCAACAGGCGCTTCGAACATCTGGAAGGATTCGCAATGCAATGACTCT	205940
Qy	5527	ACGAATTTACGGAGCTTCGATGACAGCTGGTTACTGCTATACGTCGCGGGCGTGG	5586
Db	205941	GTGAGCCTTTACAGAACGTTTGGATGATCTCTGAGATATTTGCTACCAATGTCGCGGGAGTAG	206000
Qy	5587	TCGGTTTGTATGATGGCGCGCTAAATGGGCGTGGCGCAAGCGGTGCTCGATCAACGCT	5646

Db	206001	TTGGGCTGATGATGGCAAGGGTTATGGGGTTCCGAGCGAGTGCTGATCGGGCT	206060
Qy	5647	GCATTTAGGACTGGCTTCAGCTCACTAACTATGGCGCGCACATTTAGAAGATCGG	5706
Db	206061	CGATCTAGGAATTTGCTTTTTCAGTTAACTAACTATGCTCGGGATATATTAGGATGCGA	206120
Qy	5707	AAATGGTTCGCTGCTATCTGCGCAATCTGCTGCTGATCAGCGGGATTCAGCGCGGATA	5766
Db	206121	AGGCGGGCGTGTGATCTTCCGCTAGAAATGGCTTTCACAAAGAGGGCTAATGCCAGACA	206180
Qy	5767	CGCTGACTGCACCGCAACATCGTCGAGCGCTTCGCTCACTGGCAGCGGCTTTAGTGGGG	5826
Db	206181	CACTTATTTATACCGAAACCGTCTCGCTAGCTCGGGTGGCATCCCGATTGATTGG	206240
Qy	5827	AGGCGGAACCCCTATATACATTCGGCGGATCCGGTTTACCGGGTTTACCGCTGCGCTCGG	5886
Db	206241	AAGCGGAATCTTATATACCTCCGCTCTGACAGGGCTTTGTTGGTTTACCTTTGCGGTT	206300
Qy	5887	CGTCGGGCATCGCTACGCTCGCGCGTTCATTCGCGAAATTTGGCGTCAAAGTTTCAGCACG	5946
Db	206301	CTCGGTCATCGCTCTGCTCATGGTATCTATCTGTGAATCGGCATTAAGTGCACCAAG	206360
Qy	5947	CCGGTGTGCACCGCTGGGATTCACGGCAGCGCACAGTAAAGGTGAAAACTGGCGCTGCG	6006
Db	206361	CAGGATACGGGCTCGGGAATGCAGACAAAGAACTAATCTGTGAGAGGAAGGTACCTTTAT	206420
Qy	6007	TGTTGAAGGGCGAGGTTTGGCGATCACTTCGCTGTGCTCGTCTGAACCGCGTCGCG	6066
Db	206421	TAATGGCTGGCGCAATGAAGCGCTTATATCTCGGTGGCGATTAACACGCCCTCGGATC	206480
Qy	6067	CTGGTCTGTGGCAGCGTCTCTGTTGA	6092
Db	206481	CCAAATGTGGCAGCGACTCATTA	206506
RESULT 9			
ERWCRTIB			
LOCUS			
DEFINITION	ERWCRTIB	2415 bp	DNA linear BCT 11-APR-2001
ACCESSION	M38423		Pantoea agglomerans phytoene dehydrogenase (crtI) and prephytoene
VERSION	M38423.1	GI:148401	pyrophosphate synthase (crtB) genes, complete cds.
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
COMMENT			
FEATURES			
source			
gene			
CDS			

```
/translation="MKKTVVIGAGPGGLALAIRLOAAGIPTVLLRQDKPGGRAYVMH
DGFPTDAGPTVITDPTALEALFTLAGRMEDYVRLLEPVKPFYRLCWSESKTLDIAND
SABLEAQITQFNPRDVEGYRFLYSAQVDFQGYLRLGSLVPELFRDMLRAGPQLLKL
QAWQSVYQSVRFIDEHLRQAFSPHLSLVGNPFRTSSIYTLIHALREWGWFPFEG
GTGALVNGWMLKFTDLGGIEELNARVELVADNRVSOVRLADGRI FOTDAVASNADV
VNTYKLLSHLPVQGRAAALRKSKMSNSLFLVYFLNQHPSQLAHHITCFGRPREL
IDEITGSLHADDFSLYLHSPCVTPDPSLAPGCSAPYVLAPVPHFIGNAPLDWAQSPK
LDRIPDYLEERYMGLRSLQVLTQRIPTPADPHDILDALHLSAFSIBELLLTQSAWFRP
HNRSDDIANLYLVGAGTHPGAGIPGVVASAKATASLMIEDUQ"
1486..2415
/gene="crtB"
1486..2415
/gene="crtB"
/feature="putative"
/codon_start=1
/transl_table=11
/product="prephoene pyrophosphate synthase"
/protein_id="AAA24821.1"
/db_xref="GI:148403"
/translation="MSQPLLDHATQTMANGSKSPATAAKLFDPAIRBSVLMLYTWCR
HCCDDVIDDQTHGFSAEAAEAEATQRLRLTLTAAFEAGSMQDPFAAFQEOVALTH
GITPRMLDHLGDFAMDVAQRTYVTFEDTLRYCYHVAGVGLMMARVMGRDRLDR
ACDGLAFOLTIARDIIDDAIDRCYLPAEWLODAGLTPENYAARENRAALRAVAER
LIDAAEPYVYSSOAGLHLDLPPRCANAIATARSVYREIGIKVKKAAGSSAWDRQHTSKG
EXIAMMAAPGVIRAKTRVTPRAGLWQRPV"
```

ORIGIN

```
Query Match 17.0%; Score 1189; DB 1; Length 2415;
Best Local Similarity 68.6%; Pred. No. 1.6e-239;
Matches 1655; Conservative 0; Mismatches 755; Indels 3; Gaps 1;

QY 3678 AGAAAAATGAACGCACTTATGTGATTGGCGAGGCTTTGGCGGCTGGCGCGTGGCGAT 3737
DB 4 ATAAAGGATGAAAAAACCGTTGTGATTGGCGAGGCTTTGGTGGCTGGCGCTGGCGAT 63

QY 3738 TCGCCTGCAAGCGGGGCGATACCAACCTTACTCGAGCAGCGGACAAACCGGGCG 3797
DB 64 TCGCCTGCAAGCGGGGCGAGGATCCCAACCTTACTGCTGGAGCAGCGGGAACAAGCCGCGG 123

QY 3798 AGCGGCTATGTGTTTGAGCAGAGTGGCTTTTACCTTCGATGCGGACCCACCGGTGATCAC 3857
DB 124 TCGGGCTACGCTTGGCATGACCGGCTTTACCTTTGAGCGCGGGCGGCGGTGATCAC 183

QY 3858 CGATCCAGCGGCATCGAAGAGTGTGTCAAGCTGGCAGGAAATCGCTCAGCGATTACGT 3917
DB 184 CGATCCAGCGGCTTGGAGCGGTGTTCACCTCGCGCGGCGGCGATGAGGATTTACGT 243

QY 3918 CGAGCTGTGCGGTAAACCCCTTCTATCGCCTGTGCTGGGAGATGGAACAGCTTGA 3977
DB 244 CAGGCTGTGCGCGTAAACCCCTTCTACCGACTCTGCTGGGAGTCCGGGAGACCCCTCGA 303

QY 3978 TTACGACAAATACAGCCGCTGCTGGAGCAGCAGATCGCCACGTTCAATCCGCAAGATGT 4037
DB 304 CTATGCTAACGACAGCGCGAGCTTGGAGCGCAGATTACCCAGTTCAACCCCGGAGCT 363

QY 4038 AGAAGGCTATCGTCAATTTCTGCTTATTCAGTGAAGTATTTAGAGAGGTTATCTGAA 4097
DB 364 CGAGGCTACCGGCGCTTTCTGCTTACTCCAGCGGTATTTCCAGGAGGATATTTGCG 423

QY 4098 ACTCGGACGCTGCGGTTCTGCAAGTGTGTCAGTCTGCGGCGGCGCGCGGCTGGG 4157
DB 424 CCTCGGACGCTGCGGTTCTCTCTTTTCGCAACATGCTGCGGCGGCGGCGGAGCTGCT 483

QY 4158 AGCTCTGCAAGATGCGCAGGCTTACAGCATGTGGCGGAAATTTATTCAGGACGATCA 4217
DB 484 TAAGCTCAGGCGTGGCAGGCTTACCAAGTCTGCTGCGGTTTTCGCGTTTATTTAGAGATGAGCA 543

QY 4218 TCTGCTCAGGCGTTTCTTCCACTCATTTGCTGCTGGCGGTAATTCCTTTTTCGAAAGTC 4277
DB 544 TCTGCGGACGCGCTTCTGCTTCCACTTCCCTGCTGCTGAGGCGCAACCCCTTCCACACCTC 603

QY 4278 ATCGATCTATCTTAATTCATGCGCTGAGCGTGAATGGGCGGTGTTTTCGCGCGG 4337
```

```
DB 604 GTCCATCTACACCTGTATCCACGCCCTTGAGCGGAGTGGGGGTCTGTTCCCTGAGGG 663
QY 4338 CGGCACCGCGCGCTGGTGTGAGGCGATGCGCGCATGTGTCGAGCATTTGGGCGGCGAGCT 4397
DB 664 CGGCACCGCGCGCTGGTGTGAGGCGATGCGCGCATGTGTCGAGCATTTGGGCGGCGAGAT 723
QY 4398 GTTACTGTAGTCCGAAGTGTAGCCACTGGAACACGAGCGGCAATCGCATTTAGCGGCGTTCA 4457
DB 724 CGAATCTCAACGCCCGGGTGCAGAGAGCTGGTGGTGGCCGATAACCGGTGAAGCAGGTCCG 783
QY 4458 GTTAGAGGCGGACGACGCTTCGATGCGCGCGCTGTGGGCTTCAATTCGCCAGCTGTGTGCA 4517
DB 784 GCTGGCGGATGCTCGGATCTTTGACACCGACCGCTAGCTCGAACGCTGACGTGTGTGA 843
QY 4518 TACCTACGACAACTGCTTGGCCACATCGCGCTGGCAATGAAACGTGGCAATTCGCTGAA 4577
DB 844 CACCTATATAAAGCTGTCTCGGCCACCATCGTGGGGCAGAAAGCGGCGGCGAGCTGGA 903
QY 4578 GCGTAAGCGCATGAGCAACTCGCTGTGTACTCTATTTTGGCCCTGAATCAGCGGCATCA 4637
DB 904 GCGCAAGAGCATGAGCAACTCGCTGTGTGTCTTACTTTGGGCTTGAACACGCTCATTT 963
QY 4638 ACAGCTCGCGCACCAACACGCTGTGTTTGGCGCGGCTTATCGTGAGTTGATCGATGAGAT 4697
DB 964 CAGCTGCGGCACCATACCATCTGTTTGGTCCCCTACCGGGAGCTGATCGACGAGAT 1023
QY 4698 TTTCAACAGCAGCAGCTGGCAGACGATTTTTCATCTTACCTGCAACGCGCCCTGCAAGCAG 4757
DB 1024 CTTTACCGGACGCGCTGGGAGATGATCTTCTCGCTCTACCTGCACTCGCCCTGCGTGAC 1083
QY 4758 CGATCGCTGCTGGCACCGCGGCTGGGCGAGCTTTTATGTTAGCGCGGTCGCGCA 4817
DB 1084 CGATCCCTCGCTCGCGGCTTCCCGGCTGGCGAGCTTCTACGTGCTGGCGCCGCGTGGCGCA 1143
QY 4818 TCTCGGCACCGCTGACATCGACTGCGCAACAGCAAGGACCGCGCTTGGCGGATCGAATTTT 4877
DB 1144 TCTTGGCAACCGCGCTGGAGTGGGCGCAGAGGGGCGGAGCTGCGGCAACGCGATCTT 1203
QY 4878 TGTCTTATCTGGAGCAGCACTACATCCCGGATTTAGCTCAGCAATTTAGTGACACACAGAT 4937
DB 1204 TGACTACCTTGAAGAGCGCTATATGCCGCGCTGCGTAGCTAGCTGCTGACCGAGGAT 1263
QY 4938 GTTTACGCGTGTGATTTTGGCGACACCGCTGATGCCCATCACGCTCGGCGCTTTTCGCT 4997
DB 1264 CTTTACCGCGCAGACTTTCACGACACGCTGGATCGCATCTTGGGATCGGCTTCTCCAT 1323
QY 4998 GGAGCGGATTTTGAAGCAAGCGCTGTTCCGCGCGCATAAACCGGATCGCGATATCAG 5057
DB 1324 CGAGCGGCTGTGACCCCAAGGCGCTGTTCCGCGCGCACACCGGACGAGCAGCATTCG 1383
QY 5058 CAATCTCTATCTGGTGGTGGCTGATCGCATCCAGCGCGGCGGCTGCGCGGCTGATCGG 5117
DB 1384 CAACTCTTACCTGGTGGCGGAGTACTCACCTGGGGCGGCAATCTCTGCGGTAGTGGC 1443
QY 5118 TTCGCGCAAGGCGCACCGCAGCTGATGCTGGAGGATCGCGCGGATGATCGACAGCT 5177
DB 1444 CTCGCGCAAGGCGCACCGCAGCTGATGATGAGGATC---TGAATGAGCAACCGCGG 1500
QY 5178 TTTACTTGAAGTAGTAACCAACCATGCGGTGGCTCGAAGAGTTTTCGCCACCGCGCGC 5237
DB 1501 CTGCTTGTACCGCCACGACGACCATGGCCAAAGCTCGCAAGGTTTTCGCCCGCTGGC 1560
QY 5238 AAGCTGTTTGTATGACCGACGCGCGCAGCAGCTGATGCTGATGCTGCTGCTGCTGCTGCT 5297
DB 1561 AAGCTGTTTGTATGATGGGCAACCGCTGGCGGAGGCGGCGCAGCAGCATGCGCTGCA 1620
QY 5298 TCGCATGATGATGATGATGGGCAACCGCTGGCGGAGGCGGCGCAGCAGCATGCGCTGCA 5357
DB 1621 TCGCATGATGATGATGATGGGCAACCGCTGGCGGAGGCGGCGCAGCAGCATGCGCTGCA 1680
QY 5358 GACGCGCAGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5417
DB 1681 GAGGCGCACCGCGCTGCGCGGCTGCGCACGCTGCGCGGCTGCGCGGCTGCGCGGCTGCGCGG 1740
```



Db 293 AGCTGGTCTGCCACGACAGCGATCTCACTGCGCGCAGCGCGCTTCTCGCTCCCTTC 352  
Qy 2717 ATTACGCTGGTGTGTCAGGTTATCAGGTGCGGTTTTCCTGCGCTGCGCGCAATCTGGAC 2776  
Db 353 CAGAGCTATCTCTGGGCGGGACAGGCGTGCATTTTCCCGCTTTTTCGGGCAACGCTCAAA 412  
Qy 2777 GGGGATATTGTTCCATCCATCAGCGATTTTGGCCGCCATCTTTACGGCGGATGGGT 2836  
Db 413 GGGGGTACGCCACCAATTTCTCCGAGCGCATGCGCGAGGTGATGAATCAGCGCTGTGC 472  
Qy 2837 GACGATCTGTGACAAACACAGCCGTACACAGGTAAACCCACGACGAGTGCAGCTGGCG 2896  
Db 473 GCGCCATCGCACCAATGCCCGGTGCGCCATGTGGCGCCGACCACTGGTGTGGA 532  
Qy 2897 GATGGCGGTAACTTGCTCGCAAGTGTGATGATGATGCTGCGGCTGCGCGACGCCA 2956  
Db 533 GCGCGGAGCGCATCAGACCGCGCGCTGCTGGAGCGGCGGCCCACTCGCTCGCGC 592  
Qy 2957 CATCTGACGCTGGGTTATCAGGTGTTTCTTGGACAGAGTGGCAGCTGGCGGACGCCAC 3016  
Db 593 CATCTGACCTCGCTACGACCTTCTCTGGGACAGAACTGGGTATGTCTCGGCCCAT 652  
Qy 3017 GGCCTGACGACGCCATCTGATGATGCCACCGTGCATCAGCAAGCGGTTATCGTTT 3076  
Db 653 GGCCTACCGGCCCATCATGATGCGCGGTGAGCAGCTGGCGGTTACCGCTTC 712  
Qy 3077 GTCTACGCTGCGCTGAGCGCGATCGGCTATTGATGAAGATACCATTAAGTTAAC 3136  
Db 713 GTCTATGTGCTGCGCTGACGATGACAGCTGCTGTTGGAAGACACCTATTATGCCGAC 772  
Qy 3137 CAGCCGCGCTGGCGAGAACCGCTGCTCAGCATCGCGCATATGCCATCAGCA 3196  
Db 773 GGCCTGGATTTCTCCGCGCATGCTTGGCGGCGCATTTCCGCTATGCCGCGCGCAG 832  
Qy 3197 GCGTGAACGCTGATGACGCTGCTGCTGAAGACGCGCATATTACCGATTACCTTGAGC 3256  
Db 833 GCGTGGCGCTGGATTATGTGTGCGGAGGAGGACGCGCATTTGCCGATAGGTTGGC 892  
Qy 3257 GGCACATCGATGATTTGCGAACAGCAGCAGC---GCGGCAAGCGTGCAGCGGCTGCGC 3313  
Db 893 GCGCATCAAGCGCTTTCTTCCGAGAGCGCTTCCGGGCTGCGCGCCGCGCACTGCGG 952  
Qy 3314 GCGGCGCTTTTCATGCCACACCGGTTACTCTTCCGCTGCGCGTGGCGGTAGCGGAG 3373  
Db 953 CAGAGCTGTTCACCCACACCGGCTACAGCTGCCGACGCGCATGCGCCCTGCGCGAC 1012  
Qy 3374 TTGGTAGCAGCGCTGTGCGCCACCGATGCCCTCAGCTCAGCCAAATATCGAACGCTTT 3433  
Db 1013 TCGGTGAGCGGCTTCCGACCTTTCCGGCCCGCGCTTTCGGCGCAGTGGCGCCAT 1072  
Qy 3434 GCGCGTACAGTGGCGGAAACAGCATTTTTCGCTGCTAAACCGCATGCTGTTTTG 3493  
Db 1073 GCGGCGCGGTGGAACGCGCGCGCTTCTTCCGCTGCTCAACCGCATGCTGTTCCGC 1132  
Qy 3494 GCGGTAAAGCGCAGCGCTGGCGGTGATGCAAGCTTTTACCGCTCGATGCGGG 3553  
Db 1133 GTGCGCGATCCGAGCGGCGCTATGCGCATTTCTCAGCGCTTCTACGGGCTGTGGAAGAC 1192  
Qy 3554 TTAAATAGCGCTTTTACGCGGGCAACTGCGCTGCGCATAAACCGCGATTCTGTGC 3613  
Db 1193 CTGATCGCGGTTTCTACGCGCATCGCTCACCTTCCGACAGGCGGCGCATCTTTTCG 1252  
Qy 3614 GGAAGCGCGGTGCCATCGGTGAAGCGCTGCGCGCTGTGAAATTTCTGTCGAACCA 3673  
Db 1253 GCGCGGCCACCGGTATCGGTGTTTTCAGGGGCTCTCTGCTGCTTGTGAGACGAAGCGCA 1312  
Qy 3674 GGAAGAAAAATGAAA-----CGCACTTATGTGATGGCG 3709  
Db 1313 CCGGGTTCCCATGATGCTGATGCGAACCAACCATTCGCGCGCGCGCTCATCGGT 1372  
Qy 3710 CAGGCTTTGGCGGCTGGCGATTGCGCTGCAAGCGGCGGCGCATACCAACCACT 3769  
Db 1373 CCGGCTTCCGGCGGCTCTCTCTCGCCATCGGCTGACGGCGGCGCATCCGACACCG 1432

Qy 3770 TACTCAGCAGCGCGCAAAAACCGGCGGAGCGGCGCTATGTGTTTGAAGACAGTGGGTTTA 3829  
Db 1433 TGTTCAGCAGCGGAGCAAGCCCGCGCGCTACGTTCTATGACGAGGACGGGTTCA 1492  
Qy 3830 CTTTCGATGCGGACCCACGCGTGATCAGCATCCAGGCCCATCGAAGAGTTGTTCAACGC 3889  
Db 1493 CTTTCGATGCGGCGCCACCGTCAACCGATCCCACTGCTGGAAGAGTGTTCGCG 1552  
Qy 3890 TGGCAGGAAAATTCGCTCAGCGATTACGTCGAGCTGATCGCGTAAGCCCTTCTATCGCC 3949  
Db 1553 CCGCGCGCGCAAGCTGTCCGACTATGTGAGCTGATGCGGCTCTCGCGCTTCTACCGGC 1612  
Qy 3950 TGTGCTGGAGAGATGGCAAAACAGCTTGATTACGAATAATCAGCGCTCTGCGGACGAC 4009  
Db 1613 TGTCTGCGCGACGCGCGCGCAGTTTCGACTATGTGAACACAGCGCGCGCTGAGCGCC 1672  
Qy 4010 AGATCGCACGTTCAATTCGCAAGATGTAGAAGGCTATCGTCAATTTTCTTGCTATTCAC 4069  
Db 1673 AGATCGCACCTTCGATAAGCGGAGCTGGAGGCTATCGCGCTTCTCGCTATTCG 1732  
Qy 4070 GTGAAGTATTTAGAGAGGTTATCTGAACCTCGGACCGGTGCGGTTTCTGCAAGTCCGTG 4129  
Db 1733 AGCGGCTTTCGAGGAGGTTATCTCAAGCTCGCGCGGTGCGGTTCTCGAATTCCTCCA 1792  
Qy 4130 ACATGCTGGCGTCCGCGCGAGTTGGAGCTGTGAAGCATGGCGCAGCGTCTACAGCA 4189  
Db 1793 GCATGATGAAGCGCGCGCCAAAGCTGGTGAAGCTGGAAGCTGCGGCTCGGCTCTATGCCA 1852  
Qy 4190 TGGTGGCGAAATTTTATCAGGACGATCATCTCGCTCAGCGGTTTCTCTTCCACTCATTCG 4249  
Db 1853 TGGTCTCGGCTTCACTCGGACGATCACCTCGGCGGCTTCTGTTTCAATTCCTCTG 1912  
Qy 4250 TGGTGGCGGTTAATCTTTTGAACGCTCATCGATCTATACCTTAATTCATGCGCTGGAGC 4309  
Db 1913 TGGTGGCGGCAATCCCTTCTCCACTCTGTCATCTATGCCCTCATCTCATGCGTGGAGC 1972  
Qy 4310 GTGAATGGGCGGTGTTTTCGCGCGGCGGACCGCGCGCTGTGTGAGGCGATGCGCG 4369  
Db 1973 GCAAGTGGGCGGTGTTTCTTCCCGCGGGGAAACCGCGCGCTTGTGTGCGCGCATGGTCA 2032  
Qy 4370 GACTGTTGAGGACTTGGCGCGGAGCTGTGTAATCTGAATGCCGAAGTGAAGCTGGAAA 4429  
Db 2033 AGCTGTTTACCGACCTCGCGCGCGAGATCCGCTCACAGCCGCGTGGAGAGATCGTGG 2092  
Qy 4430 CCAGCGGCAATCGAATTAGCGCGCTTCAAGTGAAGCGGCGGACGCTTCGATGCGCGCG 4489  
Db 2093 TGGAGGCGCAGCGCGCCACCGCGGTGAAGCTCAAGAGCGCGAGACCTTCCGCTTCGACC 2152  
Qy 4490 CTGTGCGCTCCTAATGCGGAGCTGTGTGCTATACCTACGAATACTGCTTCCGACCATCCGC 4549  
Db 2153 TCGTGCCTTCAATGCGGACGTGGTGCACACCTATCGCCACCTTCTGCGCGGCGCGCGC 2212  
Qy 4550 TGGCAATGAACGTCGCGCATCGCTGAAGCGTAAAGCGCATGAGCAACTCGCTGTTTGTATC 4609  
Db 2213 GCGCTGCTCGGAAGCGCGCGCTGGCCAAAGCGCGCATCTCATGTGCTGTTGCTCA 2272  
Qy 4610 TCTATTTTGGCTGAAATCAGCGCGCATGAACAGCTGCGCACCAACCGCTGTGTTTGGCC 4669  
Db 2273 CCTATTTGGGCGCGCGGCACTTGGACCATCTCCAGCACCAACAGCGTGTGTTGCTC 2332  
Qy 4670 CCGGTTATCGTAGTTGATCGATGAGATTTTCAACGAGCAGCAGCTGGGACAGATTTT 4729  
Db 2333 CCGCGTATCGGGCGCTCGTGGACGAGATCTTCAAGGGCCGAACTTCCCGACGACTTCT 2392  
Qy 4730 CACTTTTACCTGACCGCGCTCGCAGCAGGATCTGCTGCTGGCACCGCGCGCTGGCGCA 4789  
Db 2393 CGCTCTATCTCACGCGCCACCGCTGACAGCAAAATCGTGGCGCGGAGGCTGTACCG 2452  
Qy 4790 GCTTTTATGTGTAGCGCGGTGCGCATCTCGGCAACCGCTGACATCGATGCGCAACAG 4849  
Db 2453 CTTTCTAGTGTCTCTCGCGGTGCGCATCTGGGCAAGCGGACATCGACTGGAGGTGG 2512



Qy	4850	AAGAACCGCGCTTGGCGGATCGAATTTTTCGTTATCTGGAGACGACACTACATGCCGGGAT	4900
Db	2513	AAGCCCGCTCTATTCGGGACCGCATCTCGCCCATCTGGAAGACGGCTACTCCCGGGGC	2572
Qy	4910	TAGTCAGCAATTAGTGACACACAGATGTTTACGCGTTTGTATTTTCGGGACACGCTGC	4969
Db	2573	TGGCGGACAGCCTCGTCACTTCGGGCATCTCACCCCTTCGGCTTCGGCGACGAATCTGT	2632
Qy	4970	ATGCCCATCAGGCTCGGCGCTTTTCGCTGGAGCGGATTTTTCGACGCAAGCGCTGTGTTC	5029
Db	2633	CCGCCCATCAGGCTCGGCGCTTCTCGGTAGAACCGCTCCCTCACCCAGTCAGCGTGTGTTC	2692
Qy	5030	GCCCGCAATACCGCGGATGCCGATACAGCAATCTCTATCTGGTGGGTGCGGTACGCATC	5089
Db	2693	GGCGGCACAAACCGGATGCCAAGATTGCCAACTCTATTTTCGCGGGGCGGACCCATC	2752
Qy	5090	CAGCGCGGGGTGCCCGGGTGTATCGGTTTCGGCCAAAGGCCACCGCCAGAGCTCATGTCTG	5149
Db	2753	CGGGGCGGGGTGCCCGGGGTGTGGGTCTCGGCCAAGGCCACGGCGGGGCTCATCTCTCG	2812
Qy	5150	AGGATCGGCGC-----GAATGAATCGACAGCCTTTTACTTGAGCAAGTAAC	5194
Db	2813	CGGATCTGGGCGTGACCGCGCGATGAGCAGCCTTCGCGCTACCGTCTCGTGCCACGGA	2872
Qy	5195	GCAACCATGSCGCTGGGCTCGAAGAGTTTCGCCACCGCGCCCAAGCTGTTTCATGCACC	5254
Db	2873	GGAGGCATTAACAAGGCTCCAAAGACTTCGCCCGCGCGCGGCTGTTTCGGCCGCG	2932
Qy	5255	GACGCGCGCAGCACGCTGATGCTGTATGCGTGTGTGTCATCTGCATGATGTGATGA	5314
Db	2933	CATGCGGGAAGACGGGTGATGCTCTATGCTGTGCGGCCCATCTGCACAGACGTGTGTGA	2992
Qy	5315	TGGGCAACGCTGGGCGAAGCGGCACGCAGCATGCGGTTCGAAGACGGCGAGGACGCTAT	5374
Db	2993	CGGGCAGGAATCGGGGCACGGGCGGTCTCCAGCGCGGCTTCGCGCCCGCAG---CGGCT	3049
Qy	5375	GCAGCATCTGCAAAATTGAAACCCCGCGGCTACAGCGGCGCGCACATGGATGAACCGGC	5434
Db	3050	CGATGGCTCTATAGGAGAGACCGGCACACGCTATCGCGGCGCGCCTCCGCGCATCCGC	3109
Qy	5435	GTTTAGGGCGTTTACGGAAGTGGCGATCATTCACAGCTGCGGCACAACTGGGCGTTTGA	5494
Db	3110	CTTCGCGCGCTTCGCGGAGTGGTGAAGCGCAACGACATTCGCGAGCGGTATCCCTTCA	3169
Qy	5495	TCATCTGGAAGCTTCGCTATGGATGCACGCAAGCAATTCACGCGAGCTTCGATGACAC	5554
Db	3170	ACTGCTGAGGGCTTCGCGATGACGTTGGAAGCGCGCGCTACGACACCTCGATGACAC	3229
Qy	5555	GCTGCGTTACTGCTATCATCGTCGCGGGCGTGGTCGTTTGTATGATGGCGCGCTAATGGG	5614
Db	3230	GCTCACCTATTGCTACCACTGCGCGCGGTGGTGGGCGTGTATGATGGCCCTCATCATGG	3289
Qy	5615	CGTCGCGCACGAAGCGGTGCTCGATACGCTCGGATTTAGGACTGGGCGTTCAGCTCAC	5674
Db	3290	CGCGCGGATGAATGTGTCTGACACCGGCGAGCGACCTCGGCTTCGCTTCAGACTCAC	3349
Qy	5675	TAACTTCGCGCGCATTTGTAGAAGATGCCGAAATGGTTCGTGCTATCTGCGCGCAATC	5734
Db	3350	CAACATCGCCCGCAGCTGATGGAGGATGCGCCCATCGCGCGCATCTACCTTCGCGCGGC	3409
Qy	5735	CTGGCTCGATCAGGCGGGATTAACGCGCGGATACGCTGACTGCACCGCAACATCTGTGACG	5794
Db	3410	CTTCTCGAGAGGCGCGGTGCGGAGGCGGCATCGCCGCGCGGAGCACCGCGCGGC	3469
Qy	5795	GCTCGGCTCATCTGGCAGCGCGTTTATGTCGCGAGGCGGAAACCTTATTATCATCTCGGCG	5854
Db	3470	GGTGGCGCGGTGTGGCGCGCTGCTCGACGTGGCCGAGGCCATTATACGACGAGGCCCT	3529
Qy	5855	ATCCGTTTACCGGTTTACCGCTTCGCTCGGCTGGGCGGCATCGCTACGGCTCGCGGCT	5914
Db	3530	CATCGGCAATGCGCGCGCTGCGGCTTCGCGCGCGCGGCAACGCGTGGGCGCGCGGTGGGT	3589
Qy	5915	TTATCGGCAAAATTGGCGTCAAAAGTTTCAGCACCGCGGCTGTGACGCTCGGGATTACGGCA	5974

Db	3590	CTATCGCGCCATCGGCAATCGAGTTCCTAAGCGCGGCGCGCAAGGCGTGGGATACCCGCGT	3649
Qy	5975	GGCACCAGGTAAAGGTGAAAACATGGCGCTGCTGGTGAAGAGGGGCGACGTTTGGCGATCAC	6034
Db	3650	CTCCACCTCCAGCGGCCAGAGGCGCGCTTTCTCGCGGGTGGCTCGCGCAGGCGCCCTCGC	3709
Qy	6035	TTCCGGTGTGTCTCG---TCTGAACCGCTCCGGCTGCTGTGGCGAGCGTCTCTCGTTG	6091
Db	3710	CACCGGTTTCAGAGCGCGCGCGCGCGCTTCTTGGACGCGTCCCGCATG	3769
Qy	6092	A 6092	
Db	3770	A 3770	
RESULT 11			
AF218415			
LOCUS	6635 bp	DNA	linear
DEFINITION	Bradyrhizobium sp. ORS278 geranylgeranyl synthase (crtE), lycopene cyclase (crtY), phytoene desaturase (crtI), phytoene synthase (crtB), and beta-carotene ketolase (crtW) genes, complete cds.		BCT 22-JUN-2000
ACCESSION	AF218415		
VERSION	AF218415.1	GI:8650414	
KEYWORDS			
SOURCE	Bradyrhizobium sp. ORS278		
ORGANISM	Bradyrhizobium sp. ORS278		
REFERENCE	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Bradyrhizobiaceae; Bradyrhizobium.		
AUTHORS	1 (bases 1 to 6635) Hannibal,L., Lorquin,J., D'Ortoli,N.A., Garcia,N., Chaintreuil,C., Maillon-Boivin,C., Dreyfus,B. and Giraud,E.		
TITLE	Isolation and characterization of canthaxanthin biosynthesis genes from the photosynthetic bacterium Bradyrhizobium sp. strain ORS278		
JOURNAL	J. Bacteriol. 182 (13), 3850-3853 (2000)		
MEDLINE	20309720		
PubMed	10851005		
REFERENCE	2 (bases 1 to 6635) Giraud,E. and Angles d'Ortoli,N.		
AUTHORS	Direct Submission		
TITLE	Submitted (02-SEP-1999) Laboratoire des Symbioses Tropicales et		
JOURNAL	Mediterraneenes, TA/10J, Campus de Baillarguet, Montpellier 34398, France		
FEATURES	Location/Qualifiers		
source	1..6635		
	/organism="Bradyrhizobium sp. ORS278"		
	/mol_type="genomic DNA"		
	/strain="ORS278"		
	/db_xref="taxon:114615"		
	complement(273..1223)		
gene	/gene="crtE"		
CDS	complement(273..1223)		
	/gene="crtE"		
	/codon_start=1		
	/transl_table=11		
	/product="geranylgeranyl synthase"		
	/protein_id="BAF78199.1"		
	/db_xref="GI:8650415"		
	/translation="MHKPVDTDTAAFTQLDRWGRIGEVAENAFGTTVPAPLQAGMSHVLAKGGRKRYMLVIALGSLDGVPEQLLSAVAETIHAAASLVVDDLPCMDRRRSQPAHVAFGEMTAIILSIALIARAMEVAFDRQLSPARSSI VDTLSHATGPAALCGGQDLYPPVYATQDILHRYQRTSLFVAAFRCPALLAEVDEPTELRIARAQLGVAFQIFDILLDTGDHAHKGVDQDGTVTLATLLGPARAAERAADLAAVQVRELRTVGRGALDLIRMAARIAGTGGKSGRDDLRLPHAG"		
	1446..2630		
gene	/gene="crtY"		
CDS	1446..2630		
	/gene="crtY"		
	/codon_start=1		
	/transl_table=11		
	/product="lycopene cyclase"		
	/protein_id="BAF78200.1"		
	/db_xref="GI:8650416"		

/translation="MSRDADVIVIGGLAGCLIALRLTDARPDRLRVIIIEGASAGN  
HTWSFGDISDQAHWGLRVLGHRWPGFEVFAEHAIRKLSTAYLSMTSTRURAEVQ  
RPERILDATAISADHADVLEGGRTLRAPCIDAGRPPVGLAGFKELGLEVR  
LAAPGLDVPIMYDATTVAQSDGYRFTYLPDLPORLLIEDTYSGGGLPEQLVHORI  
ARYALAKQWIAEIIITRAEQDGLPVILAGPSGLVSKPSPRVRGLAALLVHETTYSL  
PDARVADLLTARLAQRGALSADARETIDGYRTIWRERGYYRFLNRMLPKAARPE  
RHRILARFVGLDQALIERFYARIQPDQKLRFVMEHMLKPPPIPSALACLPEASAFK  
TP"

gene

2627. .4147

CDS

/gene="crtI"

2627. .4147

/gene="crtI"

/codon\_start=1

/transl\_table=1

/product="phytoene desaturase"

/protein\_id="AAF78201.1"

/db\_xref="GI:8650417"

/translation="MSQEMONAKTAVVIGSGFGGLSLAIRLOSAGIATTLVEKRDKFG  
GRAYVEQDGFDFDAGPTVITDPTCLQELFALSGRKLENYVELMVPSPFYQLRWEDGA  
TFDYVNDQALERQIAAFCPADVDGTRFRPSYSERLLBEGYVKLGVPPDFRSMVRV  
APQLVALQSYRSYKSVQYVSDHQLRQAFSFLSLVGNPPFATSIIYALIHAEERW  
GWFRPGGTGLINKGLQVLFKDLGVEVTLTSVSRISVETANGRVSVAVDEGRFAAD  
IVASNAVHTYRDLKDEPLARPTAQLMRKRFMSLFIYVFLGLRRRHEPLKHHIIL  
FGRRYRELINIEFKGPALEDESILYHAPSVTDPSLAQGCSTYIVLSPVPHLAAPIL  
DWSVGPRYRDLIYLEARILPGLKSLDLCRIFTPQDFNTELNHAGSAPSLPIL  
TQSAYFRAHNADKIKGLVGLVAGTHPGAGIPGVVGSAKATARVILEQRELAQA"

gene

4144. .5151

CDS

/gene="crtB"

4144. .5151

/gene="crtB"

/codon\_start=1

/transl\_table=1

/product="phytoene synthase"

/protein\_id="AAF78202.1"

/db\_xref="GI:8650418"

/translation="WTLRDNHDLARLSEAVINGSKSPAASKLFDSTRASVHLLYA  
WCRHDDVIDGDLGIRGVGAPGQIGLQMLRDOTAQALEGAPMRDPVFGQLQVY  
QOHAIPHHVIFELLGFMVDVREYETLSETLDYCYHAGVGVYVMSAINGAREBAT  
LDRAADGLIALQLTNIARDVIEDAQTRMYPQWLCEAGVPAEVAISPPQRAVFRV  
VARLDVAEQFYEAEEAGIARLPVCAMAVETARVVYQIGREVMKRGPGAWDARIAT  
TGAQLKGAIGRSALTGLTRTWVKVPAREHNLWTRPKARRAASDAQSPANACSNRQ  
Y"

gene

5123. .5899

CDS

/gene="crtW"

5123. .5899

/gene="crtW"

/codon\_start=1

/transl\_table=1

/product="beta-carotene ketolase"

/protein\_id="AAF78203.1"

/db\_xref="GI:8650419"

/translation="MHAATAKATEFGASRRDDARORRVGLTLAAVILAAWLVHVLGM  
FEMPLTSLPALPLVLTWLYVGLFTIAHDMHGLSVLPKPKQVNRRIQGLCLFLY  
AGPSFDALNVEHKKHRRPGTADDFDEVPHGFHWPFASFLLHYFGKQVALITAAV  
SLVYQLVAVPQNLILLFWALPGLLSALQLFTFGYLPKPKATQFPADRHNRATSEFP  
AWLSLLTCTCFHGFHHEHLHPDAPWRLPEIKRRALERRD"

## ORIGIN

Query Match 15.8%; Score 1105.8; DB 1; Length 6635;

Best Local Similarity 58.4%; Pred. No. 5.6e-222;

Matches 2082; Conservative 0; Mismatches 1417; Indels 66; Gaps 6;

QY 2502 GTGCTGGCGGAGCGACCTGATGCGCACCAATACATGATGTGATTTTGGTCGTGTGGAC 2561  
DB 1426 GAGCGCGTGCCTGTGTCATGATGTCGAGATGCCGACGTCACTGTCATCGCGCGGCTC 1485  
QY 2562 TGGCGAATGGCTTGATTCGCTCGCTGCGTCAATTGCGAGCCACAACCTGAAATCGCTGT 2621  
DB 1486 TTGCGCGCTGCTGATCGCTTTGCGGCTCACCGACGACGCGCGGATCTGCGCGTCGTCA 1545  
QY 2622 TCGTGGAGCGCATCGGCATCGGCAGGCAATCATACCTGGTCGTTTCATACACGGATC 2681  
DB 1546 TCATCGAAGGCTCGGCCAGCATCGCGCGCAATCACACCTGGAGCTTCTTTGGAACCGATA 1605

2682 TCAGCGCCGAAACAATTCTGCTGGCTGCAACCCGCTGATTACCGTGGTGGTTCAGGTTATC 2741  
DB 1606 TCTCGTCCGACCCAGCAGCGCTGGCTCGAGCGGCTCGTGGTTCATCGTGGCGGGTTATG 1665  
QY 2742 AGTTCGGTTCCTCTGGCTGCGCCCAATCTGAGCGGGGATTTATGTTCCATCGCATCAG 2801  
DB 1666 AGTTCGGCTTCGCCCAACAATGCCATCCGTCCTGACCGCCCTACCTCTCATGAGCTCGA 1725  
QY 2802 GCGATTTTTCGCCGCCATCTTTACGCGCGGATGGGTGAGCATCTGTGGACAAAACACAGCCG 2861  
DB 1726 CGCGCTGCTGCGCGAGGTTCGAGCAGCGTTTTCGAGGAGGATCTGCGCGCAGCGACCG 1785  
QY 2862 TACAACAGGTAAACCCAGCAGGTGACGCTGGCCGATGGCCGTGAACTTGTGTGCGCAAG 2921  
DB 1786 CCATCTCGCAACCGCAGACCATGTCGTCTGAGAGGGGCGCCACCTTTCGCGCGCCCTC 1845  
QY 2922 TGGTATGTAGTGGTTCGCCGCCCTGACGCGACGCCACATCTGAGCTGGTGGTTCAGGTTT 2981  
DB 1846 GCGTGTATCGATCGCGCGCGCGCGCGCGGCTCGCTCTCGGGTTTCAGAACT 1905  
QY 2982 TTCTTGGACAAGAGTGGCAGCTGGCGCAGCGCAGCGCTGCGAGCAGCGATCTCTGATCG 3041  
DB 1906 TCCTCGGCTCGAGGTGGGCTGGCGCGCGCAGCGCTCGATGTGCCATCGTATCG 1965  
QY 3042 ATGCCACCGTTCGATCAGCAAGCGGTTATCGTTTGTCTACACGCTGCGCGCTCAGCGCG 3101  
DB 1966 AGCGACGCGTTCGCGCAGAGCGAGCGCTATCGCTTCGTCTACACGCTGCGCGCTCGACCC 2025  
QY 3102 ATCGCTATTTGATTAAGATACCATTCATTAGTTAACACGCCCGCTGGCGGAGAACACCG 3161  
DB 2026 AGCGCTGTGATCGAGGACCACTACTATAGCGACGCGCGGAGTTGCCGAGCAGGTGC 2085  
QY 3162 CTCGTACGACATCGCCGACTATGCCAATCAGCAAGGCTGGAGCTGAGTACGCTGCTGC 3221  
DB 2086 TGATCAGCGATCGCGGCTTACGGCTCGCCAGCGCTGGCAGATCGCCGAGATCATCC 2145  
QY 3222 GTGAAGACCGGCATATTACCGATTACCCGTAGCGCGCAACATCGATCGATTCTGCGAAC 3281  
DB 2146 GCGCGGAGCAGGCTGTCTGCGCCGCTTCTTGGCGGGCATCCGCTCGGCTGGTCAGCA 2205  
QY 3282 AGCAGCGCGGCAAGCGTGCAGCGCTGCGGCGCGGGCTGTTTCATGCCACACCGGTT 3341  
DB 2206 AGCCGACAGCGCGCGCGCGGCTGCGGCGCTTCTCGTGCATCGCAGCAGCGGCT 2265  
QY 3342 ACTCTTCGCGTTCGCGCTGCGCTAGCGGAGTTGAGCAGCGCTG-TTGCCACCGCAT 3400  
DB 2266 ATTCTGCGGACCGGTCGCGTTCGCGATCTCTGACGCGCGGCTTGGCGAACGGG 2325  
QY 3401 GCGCTTCACGCTCA-----GCCAACATATCGAACGCTTTTGGCCGTCAGCAGTGGC 3449  
DB 2326 GCGCGCTCTCGAGCGCTGACGCGCGCGAGACGATCGATGGCTACGCGCGGACGATCTGC 2385  
QY 3450 GCGAACACGATTTTTCGCTGTCTGATAACCGCATGCTGTTTGGCGCGGTGAGCGCGACG 3509  
DB 2386 GCGCGCGCGGCTACTATCGCTTCTTGAACCGGATGCTGTTTCAAGCGCGCGCGGCTCGCG 2445  
QY 3510 AGCGTGGCGGTGATGCAAGTTCGCTTTCACCGCTGCGCGGTGATTAATACCGCTTTT 3569  
DB 2446 AGCGTACCGGAATCTCGCAGCATTTTACGGTCTCGATCAGGCCCTGATCGAGCGCTTCT 2505  
QY 3570 ACGCCGCGCAACTGCGGCTTCGCGGATAAAAGCGGAT-----TCTGTGCGGCAAGC 3620  
DB 2506 AGCGCGCGCGGATCCAGCGCGGAGACAGCTGCGGCTCTTCATGCATATGCTGATGAAGC 2565  
QY 3621 CGCGGTCGCCATCGGTGAAGCGCT-----GCGCGCGCTGTTGAATT 3662  
DB 2566 GCGCGATCCGATCTCGTCCGCGCTCGCTGCGCTGCGCTGAGCGGAGCGCTTCAAGACCC 2625  
QY 3663 CTGTGCAACCGAGGAGAAAAATGAAGACATTTATGTTGGCGGCGGCTTTGGCG 3722  
DB 2626 CATGAGTCAAGAGATGCGAAACGCGCAAGCAGCAGTGTGTGATCGGTTTGGGTGG 2685

QY 3723 CCTGCGCTGCGGATTCGCTCAAGCGCGGCGCATACCAACCACTTACTCGACAGCG 3782  
Db 2686 GCTTTTCGCTCGCATACGAGCTGCAATCGCGCGGATCGCGCAACGCTCGTGGAGAGCG 2745  
QY 3793 CGACAAACGGGCGGAGCGGCTATGTTTGAAGACAGTGCGCTTACTTCGATGCCGG 3842  
Db 2746 GCACAAGTTTCGGCGCGCGGCTATGTTCTACAGACGAGGACGCGCTTCACTTCGACGCGG 2805  
QY 3843 ACCCAAGGTATCACCGATCCAGCGCGCATCAAGAGTTGTTCAAGCTGGCAGGAAATC 3902  
Db 2806 GCCCACTGTATACCGATCCGACTGCTGACGAGAGTGTTTCGGCTCTCGGGCGCAA 2865  
QY 3903 GCTACGCGATTAAGTCGAGCTGATGCGGTAAACGCTTCTATCCCTGTGTGCTGGGAAGA 3962  
Db 2866 GCTCGAGAACTATGTCGAGCTGATGCCGTCTCCCACTTCTATAGCTGCGCTGGGAGGA 2925  
QY 3963 TGGCAAAACAGCTTGATTAAGCAATATCAGCGCGTGTGAGACAGAGATGCCACGTT 4022  
Db 2926 CGCGGCCCACTTTCGACTAGCTCAAGCAACGAGCGGAGCTGGAGCGCCAGATCGCGGCGTT 2985  
QY 4023 CAATCGCAAGATGTAGAGGCTATCGTCAATTTCTTGCTATTCAAGTGAAGTATTTAG 4082  
Db 2986 CTGTCGCGGTATGTGAGCGGTATCGCGGCTTCCGTCGTATAGGAGCGCTCTCTGA 3045  
QY 4083 AGAGGTTATCTGAACCTCGGACCGGTGCGGTTCTGCAAGTGTGCTGATGCTCGCGT 4142  
Db 3046 GGAGGGTACGTCAAGCTCGGCCACGTCGCGTTCCGGATTTCGGAGCATGTCGCTGT 3105  
QY 4143 CGCGCGCAGTTGGAGCTCTGCAAGCATGGCGAGCGCTACACATGTTGGCGAAAT 4202  
Db 3106 GCGCGCGCAGCTGGGCGCTCGAGGCTATCGCAGCGCTTATCAAGAGTGTGCAATA 3165  
QY 4203 TATTGAGACGATCATCTGCTGCTGAGCGTTTCTTCCACTCATTTGCTGGGCGGTAA 4262  
Db 3166 GCTTCCGACGAGCATCTGCGGACAGCGCTTCACTTCCACTGCTGCTGCTGCGGCGAA 3225  
QY 4263 TCCTTTTGAAGCTCATCGATCTATACCTTAATCATGCGCTGGAGCGTGAATGGGCGT 4322  
Db 3226 TCGGTTGCGGACGTCCTCGATCTACGCTGATGATGCGCTGGAGCGCGCTGCGGAGT 3285  
QY 4323 GTGGTTTCGCGCGCGGACCGGCGGCTGCT---GCAGGCACTGGCGGCTGTTGA 4379  
Db 3286 CTGGTTTCGCGCGCGGACCGGCGGCTCATCAACAGGGGCTCGTCCAGCTGTTCAA 3345  
QY 4380 GGACTTGGCGCGGAGCTGTACTGAATGCCAAGTGAAGCCAGCTGGAACACGAGCGCAA 4439  
Db 3346 GGATCTCGCGCGGAGGTGACGCTGTGACAGAGCTGAGCGCGATCGAGACGCGAA 3405  
QY 4440 TCGCATTAAGCGGCTTCAAGTTAGAGGCGGAGCGAGCTTTCGATGCGCGCTGTCGCTC 4499  
Db 3406 GCGGCTCACTGCGGCTGCGCGAGGACGCGCGTGGTTCGCGCGCGACATCGTGGCCAG 3465  
QY 4500 CAATCGCGAGTGTGCTATCTACGACAACTGCTTCGCCACCATCCGCTGGCAATGAA 4559  
Db 3466 CAACGCGGACGCTGTTTACCTACCGGATCTCTGAAGGACGAGCGCGCTGGCGCGCC 3525  
QY 4560 AGCTGCGCATCGCTGAAGCGTAAGCGCATGAGCAACTCGCTGTTGTGCTACTTATTTGG 4619  
Db 3526 GACGCGGCAATTTTGTATGCGCAAGGCTTCAAGATGTCGCTGTTGCTGATCTATTCG 3585  
QY 4620 CTTGAATACGCGCATGAACAGCTCGCGCAACACCGCTCTGTTTGGCGCGCGTTATCG 4679  
Db 3586 CTTGGCGCGCGAGCATTCGAGGCTCAAGCACCACATCATTTCTGTTCCGCGCGCTACCG 3645  
QY 4680 TCAGTTGATCGATGAGATTTTCAACAGCAGCGAGCTGGCAGACGATTTTCACTTACTT 4739  
Db 3646 CGAGCTGATCAACGAGATCTTCAAGGGGCGCGCTGTCGTGAGGATTTCTCGCTGATCT 3705  
QY 4740 GCAACGCGCTTCGAGCAGGATTCGCTGCTGGCACCGCTGCTGTTTGGCGCGCGTTATCG 4799  
Db 3706 GCAACGCGCGAGGCTGACCGATCCGTCGCTCGCGCAACAGGCTTCGAGCACCATTACGT 3765  
QY 4800 GTTAGCGCGGCTGCGCGCATCTCGGCAACCGCTGACATCGATCGGCAACAGGAAGGACCGG 4859

Db 3766 GCTCTCGCGGTGCGCATCTCGCGCGCACCGATCGACTGAGCGGTGAGGCGCGCG 3825  
QY 4860 CTTGGCGCATCGAATTTTGTCTTATCTGGAGCAGCATACATGCGGGATTAAGTCAGCA 4919  
Db 3826 CTATCGGACCGCATCTCGACTATCTCGAGGCGCGCATCTTTCGCGGGGCTGAAGTCGGA 3885  
QY 4920 ATTAGTGACACACAGAAATGTTTAGCGGCTTTTTCGCGACACGCTGCAATGCCCATCA 4979  
Db 3886 CTTGCCACCTGCGCATCTTTCAGCGCGAGGATTTCAACACCGAGCTGAATGCGCATCT 3945  
QY 4980 CGGCTCGCGGTTTTCGCTGGAGCGGATTTTGAAGCAAAAGCGCTGTTTCCGCGCGCATAA 5039  
Db 3946 CGGCTCGCGGTTTTCGCTGGAGCGGATCTTCAACGAGAGCGGATTTTTCGCGCGCACAA 4005  
QY 5040 CCGGATGCCGATATCAGCAATCTCTCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 5099  
Db 4006 CGCGACACAGAGATCAAGGGGCTGTATCTCTGTCGGCGCGGACCCATCTCGGGCGCGG 4065  
QY 5100 CTTGCCCGCTGATCGGTTTCGCGCAAGCGCACCGCGCTGATGCTGGAGGATTCGCGC 5159  
Db 4066 CATTCGCGCTGCTGCGCTCGGCAAGCGCACCGCGGAGTCTCTCGAGGACCGAG 4125  
QY 5160 CGAATGAATCGACAGCTTTTACTTGAAGCAAGTAAAGCAAAACATGGCGGCTG----- 5210  
Db 4126 CGAATCGCGGAGGCGCAATGACGCTGCGGACCAATGACCATCTTGGCGCGCTGAGCGAA 4185  
QY 5211 -----GGCTCGAAGAGTTTTCGCCACCGCGCGCAAGCTGTTTGTATGACCG 5255  
Db 4186 GCGCTATCCGCAACGCTCGAGAGTTTTCGAGCGCATCCAGCTGTTTGTATTCGCGC 4245  
QY 5256 ACGGCGCGCAGCAGCTGATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5315  
Db 4246 ACCGTGCGCAGGCTGACCTGCTCTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4305  
QY 5316 GGGCAACGCTGGGCGAAGGCGGACGCGAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5375  
Db 4306 GGCAGGATCTCGGAAATTCGGCAGGCGCTGCGCGCGCTGCGCGCAGATCGGAGCTTTG 4365  
QY 5376 CAGCATCTGCAAAATGTAACCCGCGCGCTACAGCGCGCGCATCGATGTAACCGCGC 5435  
Db 4366 CAGATCTGCGCGACAGACGCGCGCTGAGGGGCGCGCATGCGCGATTCGCTG 4425  
QY 5436 TTTAGGCGGTTTTCAGGAGTGGCGATCATTCACAGCTGCGCGCAACAACTGCGGCTTGT 5495  
Db 4426 TTTCCAGGATTTGACGCGCTGCTGAGGACGACGCGATTTCCGCACTACATGCTGCTG 4485  
QY 5496 CATCTGGAAGCTTTCGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5555  
Db 4486 CTGCTCGACGCTTTCGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4545  
QY 5556 CTGCTTACTGCTATCAGCTGCGGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5615  
Db 4546 CTGAGTACTGCTATCAGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4605  
QY 5616 GTGCGCGACGAGCGGCTGCTGATCAAGCTGCGATTTAGGACTGGCGCTTCCAGTCACT 5675  
Db 4606 GCTCGGAGGAGCGAGCTGCGACGCGCGCGCGATCTCGGATCGCGCTGCGAGCTACC 4665  
QY 5676 AACATTGCGCGGACATTTGTAGAGATGCGGAAATGCTGCTGCTGCTGCTGCTGCTGCTG 5735  
Db 4666 AACATCGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4725  
QY 5736 TGGCTCGATCAGGCGGATTAAGCGCGCATACGCTGACTGACCGCGCAACATCGTCAAGCG 5795  
Db 4726 TGGCTGCGAGGCGGCTGCGCGCGCGCGGCTGCGGAGCTGCGGAGCTGCGGAGCGG 4785  
QY 5796 CTCGCTCAGCTGGCAGCGGTTTATGTCGAGGCGGAAACCTTATATCACTTCGCGCGGA 5855  
Db 4786 GTGTTCCGCTGCTGCGCGGCTGCTGATGCTGCGGAGGAGGTTTATGAGGCGGAGCGAG 4845  
QY 5856 TCCGCTTACCGGTTTACCGCTGCGCTGCGGCTGCGGCTGCGCTGCGCTGCGGCTGCGGCT 5915





GESRAVLGGIALITEAMALLAGARGASGTVRAQLVRLISRLSPQGLCAGODLHLHAA  
KNGAGEVEQDLKTGVLFPIAGLEMLAVIKEDAEQOTQIDPFOIARVPQSYDILLE  
VYGOAALGKDTGDAAPGPRGLLAVSDLKQVSRQYEAASRAQLGHWGSKRLQAPB  
TEALLERVLPYAAAR"

ORIGIN	Query Match			
	Best Local Similarity	Score	DB 1; Length	DB 2; Length
	57.4%;	14.9%;	6335;	6335;
	Matches 2059; Conservative 1;	Mismatches 1479;	Indels 51;	Gaps 8;
Qy	2456	GGTGTCCGCGGACATCGTGGAGCAGCGCTGTGCAGCAGCAAGTCGTGTGGGGAGGC	2515	
Db	1547	GGCCCGGTTCGACAGCTGAAGCAGCACTGAAGCAGCTCGGCGCTACTCGCGGCGGAGGC	1606	
Qy	2516	GACCTGATGGCAGCAGCAATACGATGATTTGGTTCGGTGTGAGTCTGGCGGAGTGGCTTG	2575	
Db	1607	GCAGAGCGCACGTACCCCATGACGTGTGTGGCAGGGGGGGCTTTGGCAACGGGCTG	1666	
Qy	2576	ATTGGCTGCGTCTGCGTCAATTCGAGCACAACCTGAATGCTGTGTGGAGAGCGAT	2635	
Db	1667	ATCGCCCTGGCTGTGGCGCGCGGCGCCGACCTGCGGGTCTGCTGTGGACCATCGG	1726	
Qy	2636	GGCAGTCCGCGAGGCAATATACCTGGTGTGTTCATCACAGCATCTCAGCGCCGAAACA	2695	
Db	1727	GGCGGACCGTCAGAAAGGCCATACCTGGTTCATGCCACGACCCCGATCTGTGCGCGCACTGG	1786	
Qy	2696	CTTGGCTGGCTCAACCGCTGATTACCGTGCCTGTGCTGAGTTTACAGTGGCTTTTCT	2755	
Db	1787	GTGGCGCGCTGAAGCCCTGCGCGCGCCAACTGGCCGACCGAGAGGTGGCTTTCCC	1846	
Qy	2756	GCGTGGCGCGCAATCTGAGCGGGGATTTGTTCCATCGCATCAGGCGATTTTGGCCGC	2815	
Db	1847	CGCCATGCCCGGGCTGSCCACCCTGACGGGTGCTGGAGCGGGCGCTGGC---G	1903	
Qy	2816	CATCTTTACCGCGGAGTGGTGAAGATCTGTGGACAACACAGCCGTACACAGGTAAAA	2875	
Db	1904	GATGCGGTGTGCGGTTCGGGCGCCGAGATCCGCTGGGACAGCAGCATCGCTTGTGGAT	1963	
Qy	2876	CCACGACAGTACGCTGCGCGATGGCGTGAACCTGTGTCGCAAGTGGTGTGATGATGT	2935	
Db	1964	GGCAGGGGGGACATCTGCTGGCGCACCGGATCGAGCCGGCTCGTCTGGAGCGG	2023	
Qy	2936	CGCGGCTGACGCGGACCATCTGACAGCTGGGTATTCAGGTGTTCTTGGACAAGAG	2995	
Db	2024	CGGGGCGGACACCGCTCCCGCATCTGACCTTGGGTTTCCAAATTTCTTGGGTTCGAG	2083	
Qy	2996	TGGCAGCTGGCAGCGCGCAGCTGAGCAGCGATCTGATGGATGCCACCGTCGAT	3055	
Db	2084	ATCGAGACGACCGCGCCACCGCGGTGCCCCCGCGGTGATCATGGACGGGACAGTCACC	2143	
Qy	3056	CAGCAAGCGGTTATCGTTTGTCTACAGCTGCGCGCTCAGCGCCGATCGGCTATTGATT	3115	
Db	2144	CAGCGGAGCGGTACGGCTTCAITTTATCTGCTGCCCTTCTTCGGACGCGATCTGATC	2203	
Qy	3116	GAAGATACCATTTACGTTAAACAGCCCGCTGGCGGAGAAACACCGCTGTGACGACATC	3175	
Db	2204	GAGGACACGCGTATTCCGATGGCGCGGATCTGGACGACGCGCTGGCGCGCGCTCC	2263	
Qy	3176	GCCGACTATGCCAATCAGCAAGCTGAGCTGAGTACGCTGCTGCTGAAGAGCAGCGC	3235	
Db	2264	CAAGACTATGCCCGCAGCAGGGGTGGACCGGGCGGAGTTCGGCGCGAA---CGGGC	2320	
Qy	3236	ATATTACCGATTACCTGAGCGGCAACATCGATTCGATTCTGG---CAACAGCAGCGCGGC	3292	
Db	2321	ATCTGCTATTGCTTGGCCCATGACGCGGGGCTTCTGGGCGGATCAACGCGAGGGG	2380	
Qy	3293	CAAGCGTGACGGCTGTGGCGCGCGGCTGTTTCATGCCACACCGCGTTACTTCTTCCG	3352	
Db	2381	CCTGTTCCCGTGGGACTGCGCGCGGGGTTCTTTACCCCGGTACACCGGCTATTGCTGCC	2440	
Qy	3353	TCCGCGCTGGCGCTAGCGGAGTTGTTAGCAGCGCTGT---TGCCCAACCGATGCCCTCAGC	3409	
Db	2441	TATGCGCGCAGGTGGCGGACGTGTGGCGGGCTTGTCCGGGCGCGCGGACCGACGCG	2500	

Qy	3410	CTCAGCCACATATCGAAACGCTTTGCCGCTCAGCAGTGGCGGAAACAGACGATTTTCCGT	3469	
Db	2501	CTGGCGCGCCATCCCGGATACCGGATCGACCGGSCACCGCGGACCGCTTCTGCGC	2560	
Qy	3470	CTGTAAACCGCATGCTGTTTTTGGCGGTAAGCGCAGCAGCGCTGGCGGTGATGCAA	3529	
Db	2561	CTGTAAACCGGATGCTGTTCCGCGCTGCGCGCCGACCGGGCTATACCTGCTGCGAG	2620	
Qy	3530	CGTTTTTACCGGCTCGATCGCGGTTAATTAGCGCTTTTACGCGCGGCAACTGCGCTG	3589	
Db	2621	CGGTTATACCGCATGCGCGATGGAATGATCGAACGTTCTATGCCGCGCGCTGAGGTTG	2680	
Qy	3590	CGCATAAAAACCGGATTTCTGTGCGCAAGCGCGGTCGCCATTCGTGTGAAGCGCTCGC	3649	
Db	2681	GCGGATCAGCTGCGCATCGTGACCGGCAAGCTCCCATTCCTTGGCAGCGGCATACGC	2740	
Qy	3650	GCGTGTGTGAATTTCTGTGCGAAACGAGGAGAAAAATG-----A	3688	
Db	2741	TGCTGCCCGAACTGCCCTGCTGAAGAAAAACGATGAACGCCCATTCGCCCCGCGCCA	2800	
Qy	3689	AACGCATTTATGTTGGCGCAGGCTTTGGCGCTGGCGCTGGCGATTCGCTGCAAG	3748	
Db	2801	AAACCGCATCGTATCGGCGCAGGCTTTGGCGGCTGGCGCTGGCGCATCGGCTGAGT	2860	
Qy	3749	CGGCGGCGCATACCAACCACTTACTCGAGCAGCGCGCAAAACCGGGCGGACGCGCTATG	3808	
Db	2861	CGGCGGCGCATCGCCACCACTTGTGCGAGGCGCGCAAGCCCGCGCGCGCGCTATG	2920	
Qy	3809	TGTTTGAAGCAGTGGCTTTTACCTTCGATGCGGACCAACGCTGATCAGCATCCAGCG	3868	
Db	2921	TCGCGCAGATCAGGCGCATGTCTTCGACGCGGCGCGACTGTCTATCACCACCCCATG	2980	
Qy	3869	CCATCGAAGAGTGTCTCACTGCGCAGGAAAAATCGCTCAGCGATTAGCTCAGCTGATGC	3928	
Db	2981	CGCTCAAGGAGCTGTGGCGCTGACCGGCGAGGACATGGCGCGGACGTGACGCTGATGC	3040	
Qy	3929	CGGTAAACCGCTTTATGCGCTGTGTGGGAAGATGGCAACACAGCTTGTATTCACATA	3988	
Db	3041	CGGTGTCCCTTCTATCGGCTGATGTGGCGCGGCGGGAAGGTCTTTCGATTACGTGAACG	3100	
Qy	3989	ATCAGCGCTGCTGCGGACGAGATCGCCACGTTCAATTCGCAAGATGTGAAGGCTATC	4048	
Db	3101	AGGCGGATCACTGGAACGCCAAATCGCCGAGTTCAACCCGCGACACTGGAAGGATACC	3160	
Qy	4049	GTCAATTTCTTGGCTTATTCACGTGAAGTATTTAGAGGGTTATCTGAAATCTCGCAGCG	4108	
Db	3161	GTGCTTCCGTGATTTACGCTCAGGAGGTGTACAGGAGGGCTACGTCAAGCTGGGACCG	3220	
Qy	4109	TGCGGTTTCTGAGGTGGTGACATGCTGCGCTGCGCGCGGCTGAGGAGCTGTGCAAG	4168	
Db	3221	TGCGCTTCTCAAGCTGGGCGCAATGTCTCAAGCGCGCGCGCGCTGATGAGCTTGAGG	3280	
Qy	4169	CATGCGCAGCGCTTACAGCATGTGTGGCAAAATTTATTCAGGACGATCATCTGCTGACGG	4228	
Db	3281	CCTACAAATTCGTTCATCGCAAGGTTCGCACTTCAAGGACCCCTATCTGCGGCGAG	3340	
Qy	4229	CGTTTCTTCCACTCATTTGCTGTGGCGGTAAATTCCTTTTGGCAAGCTCATGATCTATA	4288	
Db	3341	CGTTTTTGTATCACACGCTGCTGTGGCGGGAATCCCTTCTCGACACGCTCATCTATG	3400	
Qy	4289	CCTTAATTCATGCGCTGAGCGGTGAATGGGGGCTGTGGTTTCCGCGCGCGCACCGCGC	4348	
Db	3401	CGCTGATCAACGCGCTGAGCGCGCGCGGGGTCTGGTTTCCGCAAGCGCGCACCAACC	3460	
Qy	4349	CGCTGCTGAGCGCATGGCGGACCTGTTCGAGGACCTTGGCGCGCGAGCTGTTACTGAATG	4408	
Db	3461	AGCTGCTCGGGCATGTTCGCGCTGTTCGAACGCGCTGGCGCGCAGATGATGCTGAACG	3520	
Qy	4409	CCGAAGTGAAGCGCAGTGAAGAACCGCGGCAATCGCATTTAGCGCGCTTTCAGTTAGAGGCG	4468	
Db	3521	CCAAAGTCCCGGATCGAGACCGAGGGCGCGGACCAACGCGGCGCTCACCTTGGCGGACG	3580	





QY 3293 CAAGCGTGCAGCGGCTTGCAGCGCGGCTGTTTTCATGCGCACACCGGTTACTCTTGGCG 3352  
DB 2111 CTTGTTCCGCTGGGATGCGCGCGGGTCTTTTCATCCGGTCAACCGGCTATTGCTGCGC 2170  
QY 3353 TCCGCGTGGCGCTAGCGAGTTGGTAGCAGCGTGT---TGCCACCGATGCCCTCAGC 3409  
DB 2171 TATGCGGCACAGTGGCGGACGTTGGTGGCGGGTCTGTCCGGCGCGCCGCGCACCGACGGC 2230  
QY 3410 CTCGAGCAACATATCGAACGCTTTGCCGTCAGCATAGTGGCGGAACAGCGATTTTTTCGT 3469  
DB 2231 CTGCGCGGCCATCGCGATTACGCGATCGACCGGGCGCGCGACCGCTTCTGCGC 2290  
QY 3470 CTGCTAAACCGCATGCTGTTTTGGCCGTTAAGCCGACAGCGGCTGGCGCGTGTATGCA 3529  
DB 2291 CTTTGAACCGATGCTGTTCGCGGCTGCGCGCCGACCGCGCTATACCTTGTCTGCAG 2350  
QY 3530 CGTTTTTACCGCTCATGCCGGTTAAATTAGCCGCTTTTACGCCGGCAACTGGGCCCTG 3589  
DB 2351 CGGTTCTACCGATGCCGATGACTGATCGAACGGTTCATGCGCGCCGCTGAGCGTG 2410  
QY 3590 CCGGATAAACCGGATTCTGTGCGCAACCGCGGTCGCCATCGGTGAAGCGCTGCGC 3649  
DB 2411 GCGGATCAGTGGCATCGTACCGCGCAAGCTCCCATTTCCCTTGGCACGGCCATCCGC 2470  
QY 3650 GCGCTGTGAATCTGTGCAACGAGGGAAGAAAATGAACGCACTT----- 3697  
DB 2471 TGCCTGCCGGAACGTCCTGCTGAAGGAAACGCAATGAACGCCCATTCGCCCGCGGCCA 2530  
QY 3698 -----ATGTGATTGGCGAGGCTTTGGGGCTGGCGTGGGATTCGCTGCAAG 3748  
DB 2531 AGACCGCATCGTGATCGCGGACAGGCTTTGGCGGGCTGGCCCTTGGCCATCCGCTGCA 2590  
QY 3749 CCGCGGGCATACCAACACCTTACTCGAGCAGCGCAAAACCGGGCGGACGCGCTATG 3808  
DB 2591 CCGCGGGCATCGCACCACTGCTGAGCGCCCGGACAAAGCCCGCGCGCGGCTATG 2650  
QY 3809 TGTTTTGAGACAGTGGCTTTACTTTCGATGCGCGACCCACGGTGATCACCGATCCAGCG 3868  
DB 2651 TCTGGCACATCAGGCGCATCTCTTCGACGCGGGCCGACCGCTCATCACGACCCCGATG 2710  
QY 3869 CCATCGAAGATTGTTTACGCTGGCAGGAATAATCGCTCAGCGATTACGTGAGCTGATGC 3928  
DB 2711 CGCTGAAAGATGTGGGCGCTTACCGGGCAGACATGCGCGCGGACGTGACGCTGATC 2770  
QY 3929 CGGTAAACGCTTCTATCGCTGTGCTGGGAAAGTGGCAACAGCTTGTATTACGACAATA 3988  
DB 2771 CGGTCTCGCCCTTCTATCGGCTGATGTGCGCGGGGGAAGGTCTTCGATTACGTGAACG 2830  
QY 3989 ATCAGCGCTGTGAGGACAGATGCGCACGTTTAAATCCGCAAGATGTAGAAGGCTATC 4048  
DB 2831 AGGCGGATCAGCTGGAAACGCGAGATCGCCCAAGTTCAAACCGGACGACCTTGAGGGATAC 2890  
QY 4049 GTCAATTTCTGCTTATTCAGCTGAAGTATTTAGAGAGGTTTATCTGAAACTCGGCACGG 4108  
DB 2891 GCCGCTTCCGTATTACGCTGAGGAGGTGTACAGAGGGGCTACGTCAAGCTGGGCACCG 2950  
QY 4109 TCCGCTTCTGCAAGTGCATGCTGCGGCTGCGCGCGGAGTTGGGACGCTCTGCAAG 4168  
DB 2951 TGCCCTTCTCAAGCTGGGCGAGATGCTCAAGSCCGCGCCGCGCTGATGAAGTTGGAG 3010  
QY 4169 CATGGCGAGCGTCTACAGATGGTGGCGAAATTTATTCAGGACGATCATCTGCGTCAAG 4228  
DB 3011 CCTCAAGTGGTCCATGTCGAAGGTGCGGACCTTCAATCAAGGACCCCTTATCTGCGGCGAG 3070  
QY 4229 CGTTTTCTTCCATCTGCTGGGGGGAATTCCTTTGCAAGTCAATGATCTATA 4288  
DB 3071 CGTTTTCTGATCACAGCTGCTGGGGGGAATTCCTTCTCGACCACTGATCTATG 3130  
QY 4289 CTTTAATTCATGCGCTGGAGCGTGAAATGGGGGCTGTGGTTTTCCGCGCGGCGCACCGGCG 4348  
DB 3131 CGCTGNAACACGGCTGAGCGGCGCGGGGTCTGGTTTCGCAAGGGCGGCAACCAACC 3190

QY 4349 CGCTGGTGCAGGGCATGGCGGACCTGTTTCAGGACCTTGGCGGCGAGCTGTTACTGAATG 4408  
DB 3191 AGCTGGTTGCGGGCATGGTFCGCTGTTTCGAAACGGCTTGGCGGCGAGATGCTGCTAAACG 3250  
QY 4409 CCGAAGTAGCCAGCTGGAAACACAGCGCAATCGCATTAGCGGCGTTCAGTTAGAGGGCG 4468  
DB 3251 CNAAGGTGCGCGGATCGACACGGACGGCGCGGCGGACCGCGGTGACCCCTGGCCGAGC 3310  
QY 4469 GACGACGTTTCGATCCGCGCTGTGGCTTCAATGCGACGTTGGTGCATACCTACGACA 4528  
DB 3311 GCGCGCCTTGACCGCGACATGGTCGCGCAGCAAACGCGGACGCTGATGACCACTATCGCG 3370  
QY 4529 ACTGCTTTCGCCACCATCGCTGGCAATGAACGTCGACATGCTGGAACGCTTAAGCGCA 4588  
DB 3371 ACTCTGCGGCATATACCGCCCGCGGACAGCGCGGCGAAATCGCTGAACGCGAAGCGCT 3430  
QY 4589 TGAGCAACTCGCTGTTTGTACTCTATTTTGGCTCTGAATCAGCGCGATGAACAGCTCGCGC 4648  
DB 3431 GGTCCATGTCGCTCTTCGTGCTGCAATTTGCGGCTTGGCGAGGGGCCCAAGGACGTGGCGC 3490  
QY 4649 ACCACACCGTCTGTTTGGCCCGGTTATCGTAGTTGATCGATGAGATTTTCAACAGCA 4708  
DB 3491 ATCACACCATCTGTTTCGCGCCCGCTACAAGGAGCTGGTCAACGAGATCTTCAAGGGCG 3550  
QY 4709 GCCAGTGGCAGACGATTTTTCATCTTACCTGCAACGCGCCCTGACGACGAGTCCGTCGC 4768  
DB 3551 CGAAGCTGGCGAGGATTTCTCGCTCTATCTGCATTCGCGCTGCAACGACCGCCCGAGA 3610  
QY 4769 TGGCACCGCCCGGCTGCGGACGCTTTTATGTGTAGCGCGGCTGCGCATCTCGGCACCG 4828  
DB 3611 TGGCGCTTCCGGGATGTCACGCAATTACGCTGCGCCCGGTGCGCATCTTGGGCGCGC 3670  
QY 4829 CTGACATCGACTGGCAACAGGAAGGACCGCGCTTGGCGGATCGAATTTTGTCTTATCTG 4888  
DB 3671 CGGACATTTGTTGGCGGTTCGAGGGCGCGGCTATGCGACCGCATCTCTGGCGAGCTGG 3730  
QY 4889 AGCAGCACTACATGCGGGATTCATGTCAGCAATTAGTGACACACAGAAATGTTTACGCCGT 4948  
DB 3731 AGGAGCGGCTGATCCGAAACCTTGCAGCGCAACTGACCAACGACGCGCATCTTCAACCCGT 3790  
QY 4949 TTGATTTTTCGACACGCTGATGCCATCAGCGCTCGGCGTTCGCTGGAGCCGATTT 5008  
DB 3791 CGGATTTTCGCGACGAACTGAACGCCCATTCATGCGAGCGCTTCTCGTTCGAGCGCATCC 3850  
QY 5009 TGACCAAGCGCTGTTTCGCGCGCATAAACCGGATGCCGATATCAGCAATCTCTATC 5068  
DB 3851 TGACGCAATCCGCGTGGTTCGCGCGCAACACCGGCAAGACGATCCGCAACTTCTATC 3910  
QY 5069 TGGTGGTGGCGGTACGATCCAGGGGGGCGTCCCGCGGCTGATCGGTTCCGCCAAG 5128  
DB 3911 TGGTCCGCGCGGCAACCCATCCGGCGCGGGGATTCGCGGCGCTGTTGGGCTCGGCCAAGG 3970  
QY 5129 CCACCGCGAGGCTGATGCTGGAGGATCGCGCGCAATGAATCGACAGCCTTTTACTTGAACA 5188  
DB 3971 CCACCGCGCAGTGTGCTGCTCGACCTGGCGAGCGCATGAGCGATCTGGTCTCT---GAC 4027  
QY 5189 AGTAACGCAAAACCATGGCGGTGGGCTCGAAGAGTTTCGCCACCGCGCGCAAGCTGTTGA 5248  
DB 4028 CTCGACCGAGGGGATCAACCAAGGTCGCAAGCTTGGCAACGCGCGCAAGCTGATGCC 4087  
QY 5249 TGACCGGACCGCGCGCAGCAGCTGATGCTGTATGGTGGTGTGCTCACTGCGATGATGT 5308  
DB 4088 GCCGGGCACTCCGCGACGACACGCTGATGCTCTATGCTGGTCCCGCACGCGGATACGCT 4147  
QY 5309 GATTGATGGGCAAAACGCTGGGCGAAGGCGGACGAGCATGCCGTGGAAGACGCGCAGGC 5368  
DB 4148 GATCGACGCTCAGGCCCTG-----GGCAGCTGCCCGAGGGGGTGAACGACCGCAGGC 4201  
QY 5369 AGCTATGACGATCTGCAAAATGAAACCGCGCGCTTACAGCGCGCGCACATGATGA 5428  
DB 4202 GCGGCTGAGCGCCTGCGCGCGCACACGCTGGCGCGCTTGCAGGGGCGACGGTCCGCTGAC 4261  
QY 5429 ACCGCGCTTTAGGGCGTTTCAGGAAGTGGCGATCATTCACAGCTGCCCGCAACACTGGC 5488

```
Db 4262 CCCGCCCTTTGC-----CTGCGCGGTGCGCGCGCATGATTTTCGCGAGGCGCTGCC 4315
QY 5489 GTTTTCATCATCTGGAAGGCTTCGCTATGATGACGCAACGAAATACACGCGACTTCGA 5548
Db 4316 CATGACCTGATCGAAGGCTTTGCGATGATGTCGAGGCGCGCACTATCGACGCTGGA 4375
QY 5549 TGACACGCTGCGTTACTGATACGTCGCGGCGTGTGCTGTTTGATGATGCGCGCGGT 5608
Db 4376 CGACGCTGCTGGAATATTCCTATCAGCTCGCGGCACTGTCGCGCGTGAATGATGCGCGCGGT 4435
QY 5609 AATGCGCTGCGCGACGAGCGGTCTCGATCACCGCTCGCATTTAGGACACTGCGGCTTCCA 5668
Db 4436 GATGCGGCTGCGCGAGATCTCTGCTCGACCGTGCCTCGACCTGGGCGTGGCGTTCCA 4495
QY 5669 GCTCACTAAATGCGCGCGACATTTGTAAGATCCGGAATATGTCGCTGCTATCTGCC 5728
Db 4496 GCTGACCAACATCGCGCGCGATGATCGACGCGCGCATCGCGCGGCTGCTATCTGCC 4555
QY 5729 GCATCTGCTGCTGATCAGCGCGGATTAGCGCGCATAGCTGACGCGCGCGTGCCTGCG 5788
Db 4556 AGCGCACTGCTGCTGATCAGCGCGGCGCGCGGTGCGCGCGCGTGCCTGCG----- 4608
QY 5789 TGCAGCGCTCGCTCACTCGCAGCGCTTTAGTGGCGGAGCGCGAACCTATTATCACTC 5848
Db 4609 --CGAGCTGTACACCGTGTCTCGCGCTGCTGGATCGCGCTGAATCTATTACGCGTC 4666
QY 5849 GCGCGATCGGTTTACCGGTTTACCGTGGCTGCGCGTGGCGGCGGCGGCGGCGGCTGCG 5908
Db 4667 GCGCGGCGGCTGCTGCGCGATCTGCGCGCGGCTGCGCGTGGTCCATCGCGCGCGGCTT 4726
QY 5909 CCGCGTTTATCGGAAATGCGGCTCAAGTTCAGCAGCGCGGCTGTCAGCGCTCGGATTC 5968
Db 4727 GCGGATCTATCGCGCATCGGCTGCGGATTCGCAAGCGCGCGCGGCGGCGGCTATCGCCA 4786
QY 5969 AC CGCAGCGCACCGATTAAGGTGAAATACTGCGGCTGCTGGTGAAAGGGG 6018
Db 4787 GCGGATCAGCAGCTGCCAAGCGCGCAAGATCGGCGTCTGCGGCGATCGGG 4836

RESULT 14
FAU62808/c
LOCUS
DEFINITION
Flavobacterium ATCC21588 geranylgeranyl synthase (crtE), phytoene
synthase (crtB), phytoene desaturase (crtI), lycopene cyclase
(crtY), b-carotene hydroxylase (crtZ) genes, complete cds.
ACCESSION
U62808
VERSION
U62808.1
KEYWORDS
GI:1842241
SOURCE
Flavobacterium sp. ATCC 21588
ORGANISM
Bacteria; Bacteroidetes; Flavobacteria; Flavobacteriales;
Flavobacteriaceae; Flavobacterium.
REFERENCE
1 (bases 1 to 5188)
AUTHORS
Pasamontes,L., Hug,D., Teesier,M., Hohmann,H.P., Schierle,J. and
van Loon,A.P.
TITLE
Isolation and characterization of the carotenoid biosynthesis genes
of Flavobacterium sp. strain R1534
JOURNAL
Gene 185 (1), 35-41 (1997)
MEDLINE
97186694
PUBMED
9034310
REFERENCE
2 (bases 1 to 5188)
AUTHORS
Pasamontes,L.
TITLE
Direct Submission
JOURNAL
Submitted (03-JUL-1996) VFCB, F. Hoffmann-La Roche Ltd., Basel,
4070, Switzerland
FEATURES
location/Qualifiers
source
1. .5188
/organism="Flavobacterium sp. ATCC 21588"
/mol_type="genomic DNA"
/strain="R1534"
/db_xref="taxon:50286"
95..982
gene
```

```

/gene="crtE"
95..982
/gene="crtE"
/codon_start=1
/transl_table=11
/product="geranylgeranyl synthase"
/protein_id="AAC44848.1"
/db_xref="GI:1842242"
/translation="MTPKQFFPLRDLVEIRLAQISQGFVSVAPLGAAMSDAALSPGK
RFAVLMVAESSGVCDAVDVAACAVMVAASLI FDDMPDMDARTRRGGQPATHV
AHGEGAVLAGIALITEAMRILLEGARGATPDQARIVASMRAMGVPVGLCAGQDLH
APDKAGIERQLKTGVLFVAGLEMLSIIGLDKAEISQLMAGFQGLGRVFSYDYL
LDVIGKASTGKDTARDTAPGPKGLMAGVQMDVAQHYRASRAQLDELDELMTLFRG
GQIADLLARVLPDHIRSA"
complement(979..1890)
/gene="crtB"
complement(979..1890)
/gene="crtB"
/codon_start=1
/transl_table=11
/product="phytoene synthase"
/protein_id="AAC44849.1"
/db_xref="GI:1842243"
/translation="MTDLTATSEAAIAQGSQSFQAQAKLMPGPIREDTVMVLAWCRHA
DDVIDGVMGSAFEGSDPOARLICALRADTLAALHEDGMSPPFAALROVARHDPD
LWPDLTGEPAMDVADREYRSKDDVLEIYTHVAGVGMARVWGVQDDAVLDRACDL
GLAFQLTNIARDVIDDAIGRCYLPDLAEAGATVEGVPVSADLYSVILRLDAAEP
YASARQGLPHLPKCAWSIAAALRIYRAITGTRIQQGPEAYRQRISTSKAAKIGLLA
RGGIDAAASRLRGESIRDGLWTRPA"
complement(1887..3371)
/gene="crtI"
complement(1887..3371)
/gene="crtI"
/codon_start=1
/transl_table=11
/product="phytoene desaturase"
/protein_id="AAC44850.1"
/db_xref="GI:1842244"
/translation="MSSAAIVIGAGFGGLAIALRIQSAGIATTIVEARKDGGRAVYVN
DQGHVFDAGPTVTPDLSRELWALSGQPMERDVLLPVPFFRLTWADGSRSEFYND
DDELIQVAGSNPADVDYRFRHDAEYVREGYLGKLTPTFLKGLQMLNAPALMR
QAVRSVSMWARFTQDPLHQAQFSLTLVGGNPFSTSIYALIHALLERGGVWFAG
GTNOLVAGMARPERLGLTLNARVTRIDTEGDRATGVTLLDQRLADTVASNGDV
MHSYRDLHGTRGRTKAAIILNORWSMFLVLFHFLSKRPENLAHSVIFGPRYKGL
MSEIFNGLRPDDFSNMLHSPCTVDPSLAPGSMSTHYVLAAPVHLGRADVDEAEAPG
YAEIRFELERRAIPDLRKHILTVSRIFSPADFTSELSAHGSAFSEVPILTQSAWFRP
HNRDAIPNFYIVGAGTHPGGIFGVVGSAKATAQVMLSLAVA"
complement(3368..4516)
/gene="crtY"
complement(3368..4516)
/gene="crtY"
/codon_start=1
/transl_table=11
/product="lycopene cyclase"
/protein_id="AAC44851.1"
/db_xref="GI:1842245"
/translation="MSHDLITAGLISGALIALAVRRDRPDARIIVMLDARGSPSDQHT
WSCHTDIUSPEWLAKLSPIRGENTDQVAFPDHSRLTTCVGSIEAGALIGLQGV
LRWNTHVATDDTGTATLDGSKIEAACVIDARGAVETPHLTVGQKFVGVLETDAPH
GVERPMINDATVPQMDGYRFTYLLPFSPTRIILEDTRYSDGDDDDGALAQASLDYAA
RRGTQGMRRERGILPTALAHDAIGFWRDHAQGAVPVGLGAGLPHVTVGYSLPYAAQ
VADATAARDLTATASARRAVRWAIADRDRFLRLNMLFRGCPDPRRYLLQRFYR
LPQPLIERFVAGRLTLADRLRIVTGRPPILPSQAVRCLPERPLQERA"
complement(4513..5022)
/gene="crtZ"
complement(4513..5022)
/gene="crtZ"
/codon_start=1
/transl_table=11
/product="b-carotene hydroxylase"
/protein_id="AAC44852.1"
/db_xref="GI:1842246"
/translation="MSTWAAAILTVILTVAAMELTAYSVHRWIMHGFLGWGWHKSHHDE
```



DHDALEKNDLYGVLPFAVISIVLFAIGAMGSDLAHWLAVGVTCYGLIYYFLHDGLVHG  
RMPFRVPRKGYLRRVYQAHRMHVAHVHGRENCVSFGFIWAPSVDSLKAEKRSGLALK  
DREGADRT"

## ORIGIN

Query Match	14.8%;	Score	1035.4;	DB 1;	Length	5188;
Best Local Similarity	56.5%;	Pred. No.	3.8e-207;			
Matches 2037;	Conservative	0;	Mismatches 1536;	Indels	30;	Gaps 5;

  

QY	2486	GTCCAGCAGCAAGTCTGTCTGCGAGGCGGACCTGATCGCAGCAACGCAATACGATGTGATT	2545
DB	4558	GGGCGCGTCTGTAAGGACCGCGAAGGGCGGATCGCAATACATGAGCCATGATCTGCTG	4499
QY	2546	TTGGTCGGTCTGGACTGCGGAATGGCTTGATTTGCGTCTGCGTCTGCGTCTGCGTCAATTTGAGCCA	2605
DB	4498	ATCGGGGCGCGGGGTGTCGGTGTGCGTGTGCGCTTGGCGCTTGGCGACCGGACCG	4439
QY	2606	CAACTGMAATGCTGTGTGCTGGAGAGCGATGCGCATCGGCGAGGCAATCATACCTGGTCTG	2665
DB	4438	GATCGCGCATCGTATGCTCGACGGCGGTTCGGGCCCTCGAACAGCACTTGGTCTC	4379
QY	2666	TTTTCATCAGCGATCTCAGCGCCGAACTTTCGCTGGCTGCAACCGCTGATTACCGGTG	2725
DB	4378	TGCCAGCACAGGATCTTTTCGCCCGAAATGGCTGGCGGCTGTTCGCCCATTCGTGCGGCG	4319
QY	2726	CGTTGGTCAAGTTACAGTGTGCTTTCTGCTGCGTGGCGGCGCAATCTGACGCGGATAT	2785
DB	4318	GAATGACAGGATCAGAGGTGCGTTTCCCGACCAATTCGCGCGGCTGACGACAGGCTAT	4259
QY	2786	TGTTCCATCGCATCAGGCGATTTTCGCCCGCATCTTTACGCGCGCATGGGTGACGATCTG	2845
DB	4258	GGCTCGATCAGCGCGCGCGCTGATCGGGCTGTGCA-----GGGTGCGATCTG	4208
QY	2846	TGGCAAAACACAGCGCTCAACAGGTAAACCCACGAGGTGACGCTGGCGGATGGCCGT	2905
DB	4207	CGTGGATACGATCTCGCAGCTTGGAGATACCGCGCGACGCTGACGCGGCTCG	4148
QY	2906	GAACTTGTGCGAAGTGTGATTTGATGTGCGGCTGCGCGGCTGCGAGCGCAATCTGCGAG	2965
DB	4147	CGGATCGAGGCTGCCTGCGTGTGACGCGCGGTGTGCGGCTGCGAGCCCGCACCTGACC	4088
QY	2966	CTGGGTATCAGTGTCTTCTGACAAAGATGGCAGCTGGCGCGCGCGCGCTGCGAG	3025
DB	4087	GTGGGTTCAGAAATTCGTGGCGTCTGAGATCGAGACGACGCGCCCCCATGGCGTGGAG	4028
QY	3026	CAGCGATCTGATGATGCCACCGCTGATCAGCAAGCGGGTATCGTTTGTCTACAG	3085
DB	4027	CGCCCGATGATCATGGACCGGACCGTTCCGAGATGGACGGGTACCGCTTCTATCTG	3968
QY	3086	CTGCGCTCAGCGCGATCTGGCTATTGATTGAAGATACCCATTAAGTTAACCGCGCGG	3145
DB	3967	CTGCGCTTCACTCCACCGCATCTCTGATCGAGGATACGCGCTACAGCGACGCGCGGAT	3908
QY	3146	CTGGCGGAGAACACCGCTGTCAGCACATCGCGGACTATGCCAATCAGCAAGCTGGAGC	3205
DB	3907	CTGGAGATGGCGGCTGCGCAGGCGTCTGCTGACTATGCCCGCAGCGGGGCTGGACC	3848
QY	3206	CTGAGTACGCTGTGCTGTAAGAGCACGCGCATATTACCGATTACCTGAGCGCGCAACATC	3265
DB	3847	GGCGAGGATCGGGCGGAAAGGGGCATCTCGCCCATCGCGCTGGCCCATGACCCATA	3788
QY	3266	GATCGANTCTGGCAACAGCAGCGCGGCAAGCGTGCAGCGGCTGCGCGCGCGGTGTTTT	3325
DB	3787	GGCTTCTGCGGCGACCAACGCGAGGGGGCGGTGCGGTTGGGTGGGGCAGGGCTGTTC	3728
QY	3326	CATGCCACCGGTTACTCTTTCGCTCGCGTGGCGCTAGCGAGTTGTTAGCAGCG	3385
DB	3727	CACCCGTCACCGGATATTCTGCTGCCCTATGCGCGCAGGTCGCGGATGCCATCGCGGCG	3668
QY	3386	CTGTTGCCCAACCGATGCCCTCAGCTCAGCAACATATGAAACGCTTTGCCCGTCAAGCAG	3445
DB	3667	CGCGACCTGACGACCGCTCGCCCGTTCGCGGGTGGCGGCTGGGCCATCGATCGCGCG	3608

QY	3446	TGCGCGGACACGCGATTTTTCGCTCTGCTAAACCGCATGCTGTTTTTGGCCGCGTAAGCCG	3505
DB	3607	GATCGC---ACCGCTTCTCGCGCTGTGAACCGGATGCTGTTTCGCGGCTGCCCGCC	3551
QY	3506	CAGCAGCGCTGCGCGTGTATGCAACGTTTTTACCGGCTCGATGTCGCGGTTAATTAGCCCG	3565
DB	3550	GACCGTCTGCTATCGCTGCTGAGCGGTTCTACCGCTGCCGACGCGCTGATCGAGCGC	3491
QY	3566	TTTTACGCGGGCAACTGCGCTGCGCGATAAAAACGCGGATTTCTGTGCGGCAAGCGCGC	3625
DB	3490	TTCTATGCGGGCGCTGACATTTGCGGACCGGCTTCGATGCTCACCGGACCGCGCC	3431
QY	3626	GTGCCCCATCGGTGAAGCGTTCGCGCGCTGTTGAATCTGTGCAACCGAGGAGAAAAA	3685
DB	3430	ATTCCGCTGTGCGAGCGCGTGCCTGCCCGAACGCCCTGCTGCGAGGAGAGCA	3371
QY	3686	TGAACCGACATTTATGTGATTGGCGCAGGCTTTGGCGGCTGGCGCTGGCGATTTCGCGTGC	3745
DB	3370	TGAGTTCCGCCATCGTATCGGCGCAGGTTTCGCGGGGCTTGGCTTGCATCCGCTGC	3311
QY	3746	AAGCGCGGGCATACCAACCTTACTGAGCAGCGGCAAAAACGCGGCGGACGCGCCT	3805
DB	3310	AATCGCGGCAATCGGACCAACATGCTGAGGCGCGGACAAAGCGCGCGCGCGCCT	3251
QY	3806	ATGTGTTGAGGACAGTGGCTTTACTTCGATGCCGAGCCACCGGTGATCACCAGTCCCA	3865
DB	3250	ATGTCTGAAACGATCAGGGGCCACGCTTCGATGCAAGGCCGCGGTGCTGACCGACCCG	3191
QY	3866	GGCGCATCGAAGAGTTTTCAGCTGGGAGGAAATTCGCTCAGCGATTACGTCGAGCTGA	3925
DB	3190	ACAGCTGGGAGAGCTGTGGGCCCTCAGCGGCCAACCGATGAGCGTGCAGCTGC	3131
QY	3926	TGCGGTAAACGCGCTTCTATCGCTGTGCTGGGAAAGATGGCAAAACAGCTTGAATTACGACA	3985
DB	3130	TGCGGTCTCGCCCTTCTACCGCTGACATGGGCGGACGCGCGCAGCTTCGAATACGTA	3071
QY	3986	ATAATCAGCGCGTCTGAGCAGCAGATCGCCAGCTTCATTCGCAAGATGATAGAGGCT	4045
DB	3070	ACGACGACGAGCTGATTCGCCAGGTGCGCTCTTCAATCCGCGGATGCTGATGGCT	3011
QY	4046	ATCGTCAATTTCTGCTTATTCAGCTGAAGTATTTAGAGAGGTTATCTGAAACTCGGCA	4105
DB	3010	ATCGCGGTTTCAACATTCAGCGGAGAGGTTCTATCGGAGGAGGTTATCTGAGCTGGGA	2951
QY	4106	CGGTGCGCTTCTGAGGTGTGATGCTGCGGCTGCGCGCGCAGTGTGGACGCTGTC	4165
DB	2950	CCACGCCCTTCTGAAGCTGGGCCAGATGCTGAAACGCGCGCGCGCTGATGCGCGCTGC	2891
QY	4166	AAGCATGGGCGAGCTCTACAGCATGGTGGCGAAATTTATTCAGGACGATCATCTGGTC	4225
DB	2890	AGGCATACCGCTCGGTCCACAGCATGGTGGCGCGCTTCATCCAGGACCGCGCATCTCGCGC	2831
QY	4226	AGGCGTTTTTCTTCCACTCATTTGCTGGGCGGTAATCTTTTGAACGTCATCGATCT	4285
DB	2830	AGGCGTTTCTGTTCCACAGCTGCTGCTGGTGGCGGGAACCGTTTTGACACGCTCGATCT	2771
QY	4286	ATACCTTAATTCATCGCTGGAGCGTGAATGGGGGCTGTGGTTTCGCGCGCGGCAACCG	4345
DB	2770	ATGCGCTGATCATCGCTGGAAACGCGCGCGCGGCTGTGTTTCGCAAGGCGCGCACCA	2711
QY	4346	GGCGCTGCTGAGGCGATGCGCGCATGTTTCAGGACCTTGGCGCGGAGCTGTTACTGA	4405
DB	2710	ACCAGCTGGTTCGCGGCGATGGTTCGCCCTGTTTCAGCGCTTTCGCGGCACTGCTCTGA	2651
QY	4406	ATGCGCAAGTGAACGAGCTGGAAACCGCGGCAATCGCATTTAGCGCGTTTCAGTTAGAGG	4465
DB	2650	ATGCGCGCTGACCGGATTCGACACGAGGCGGATTCGCGCCACGCGGCTCACGCTCTGG	2591
QY	4466	CGCGACGAGCTTCGATGCCCGCTGTGGCTTCCAATGCCAAGTGGTGCATACCTACG	4525
DB	2590	ACGGCGGCGAGTTGCGCGGATACGTTGGCCAGCAACCGCGACGCTGATGACAGCTATC	2531
QY	4526	ACAAACTGCTTCCGACCATTCGCTGGGCAATGAAACGTCGACATCGCTGAACGCTAAGC	4585

Db 2530 GGCACCTGCTGGGCCATACCCGCGCGGCGACCAAGCGCGGATCTGAACCGGACG 2471  
Qy 4586 GCATGAGCAACTCGCTGTTGTTGATCTATTTTGGCTGAATCAGCGGATGAACAGCTCG 4645  
Db 2470 GCTGCTGATGCTGCTGTTGCTGCTGCTATTCGGCTCTCCAAAGCGCCCGAGAACCTGG 2411  
Qy 4646 CCACACACAGCTGTTTGGCCCGGCTTATCGTAGTTGATCGATGAGATTTTCAACA 4705  
Db 2410 CCCACACAGCGCTATCTTGGGCCCGCGCTACAAGGGGCTGTTGAACAGATCTTCAACG 2351  
Qy 4706 GCAGCCAGCTGCAGACGATTTTTCATCTTACCTGCAGCGCCCTGCAGCAGCGATCGT 4765  
Db 2350 GGCACGCTGCGGACGATTTCTCGATGATCTCAATTCGCTGCTGACCGATCCA 2291  
Qy 4766 CCGTGGACCGCCCGCTGCGGACGCTTTATGTTTGGCTGAGCGGCTGCGGATCTCGCA 4825  
Db 2290 GCTTGGCCCGGAGGGATGTCACGCAATACGCTCTTGGCCCGCTTCGCACTGGGCC 2231  
Qy 4826 CCGCTGACATCGACTGGCAACAGGAGGACCGGCTTGGCGATCGAATTTTGTCTATC 4885  
Db 2230 GCGCGATGCTGATGGAGCGAGGCCCGGCTATGCGAGCGCATCTCGAGGAAC 2171  
Qy 4886 TGGAGCAGCACTACATGCGGGAATACGTCAGCAATTAAGTGACAACAGATGTTTACGC 4945  
Db 2170 TGGAGCGCGCCATCCCGACCTCGCAAGCACCTGACCCCTCAGCCGCACTTTCAGCC 2111  
Qy 4946 GCTTGAATTTTCCGACACGCTGCAATGCCCATCAGGCTCGGCTTTCGCTGGAGCCGA 5005  
Db 2110 CCGCGAATTTACGACCGCAACTGTGCGGCCCATCAGCGCAGCGCTTCTCGTTCAGCCGA 2051  
Qy 5006 TTTTACGCAAGCGCTGTTTCCGCGCGCATAACCGCGATCGCGATATCAGCAATCTCT 5065  
Db 2050 TCTGACGCAATCCGCTGTTCCGCGCGCATAAACCGCGACCGCGATCCCGAACTTCT 1991  
Qy 5066 ATCTGCTGGGTCCGTTACGATCCAGCGCGGCGTGCCCGCGTGCATCGTTGCGCA 5125  
Db 1990 ACATCGTGGGGCGGCGACGCTATCCGGGTGCGGGCATCCCGGGTGTCTGTTGCGAGCGCA 1931  
Qy 5126 AGGCCACCGCCAGCTGATGCTGGAGATCGCGCGCAATGAATCAGACGCTTACTTGA 5185  
Db 1930 AGGCCACCGCGAGTCTGCTGTCGACCTGCGCGCTCGATGACCGATCTCAGCGGAC 1871  
Qy 5186 GCAAGTAAGCAACCATCGCTGGGCTCGAAGATTTTCCGACCGCGCCCAAGCTGT 5245  
Db 1870 TTCCGAAGCG--GCCATCGCGCAGGTTTCGCAAGCTTCGCGCAGCGCGCAAGCTGAT 1814  
Qy 5246 TGATGACCGACCGCGCGCAGCAGCTGATGCTGATGCTGCTGCTGCTCACTGCGATGA 5305  
Db 1813 GCGCCCGGATCCCGGAGGATACGGTCTATGCTCTATGCTGCTGCGAGGATCGCGATGA 1754  
Qy 5306 TGTGATTGATGGGCAAAACGCTGGGCGAAGCGGCGACGATGCTCGAAGACGCGCA 5365  
Db 1753 CGTATCGACGGCAGGTGATGGGTTCTGCCCGGAG-----GCGGGCGGCGACCCACA 1700  
Qy 5366 GGACGCTATGACAGATCTGCAAAATGAACCGCGCGGCTCAGCGGCGCGCATGA 5425  
Db 1699 GCGCGGCTGGGCGCGCTGCGCGCGACACGCTGGCGCGGCTGACAGGACGCGCCGAT 1640  
Qy 5426 TGAACCGGCTTTAGGCGGTTTCAGGAAGTGGCGATCATTTCCACGCTGCGCAACT 5485  
Db 1639 GTCGCGCGCTTCGCGCGCTGCGCGAGGTGCGCGGCGCATGATTTCCGAGCCTTTG 1580  
Qy 5486 GGGGTTTGAATCATCTGGAAGGCTTCGCTATGATGACGCAACCAATTTACCGAGCTT 5545  
Db 1579 GCCGATGGAACCTGATCGAGGTTTTCGCGATGATGCTGCGGATACCGCAATACCGAGCT 1520  
Qy 5546 CGATGACACGCTGCTTACGCTATCACGTCGCGGCGTGGTGGTTGATGATGCGCG 5605  
Db 1519 GGATGACGCTGGAATTAATTTCTACACACGCTGCGCGGGGTCGTGGGCGTGAATGCGCG 1460  
Qy 5606 CGTAATGGGCGTGGCGGCAAGCGGTGCTCGATCAGCGCTGCGAATTTAGGACTGGCGTT 5665

Db 1459 GGTATGGGCGTCAGGACGATCGGTCGTGGATCGGCGCTGCGATCTGGGCCCTTGCCTT 1400  
Qy 5666 CCAGCTCACTAAACATTCGCGCGACATTTAGAGATGCGGAAATGGTGCCTGCTATCT 5725  
Db 1399 CCAGCTCAAGACATTCGCTCGACGCTGATCGACGATGCCCATCGGCGCTGCTATCT 1340  
Qy 5726 GCGCAATCTCGCTCGATCAGCGGGATTCAGCGCGGATACGCTGACTGACCGCAACA 5785  
Db 1339 GCCTGCGACTGGCTGCCGAGGCGG-----GCGACGCTTGAAGGTTCCGCTGCC 1289  
Qy 5786 TCGTGAAGCGCTGCCCTCACTGCGAGCGGCTTTAGTGGCGGAGCGCAACCTTATATCA 5845  
Db 1288 TTGAGACGCGCTTATTCGCTCATCATCCGCTCTTGAACGCGCCGACCTTATATGC 1229  
Qy 5846 CTGCGCGGATTCGCGGTTTACCGGTTTACCGCTGCGCTCGGCGTGGGCTATCGCTACGGC 5905  
Db 1228 CTGCGCGGAGCGGCTTTCGCTATCTGCGCGCGCTGCGCGTGGTGCATGCGCGCGC 1169  
Qy 5906 TCGCGGCTTATCGCAAAATTTGCGCTCAAGTTTACGACCGCGGCTGACCGCTCGGA 5965  
Db 1168 GCTCGTATCTATCGCAATTCGCGACGCGCATCCGCGAGGTGGCCCCGAGGCTATCG 1109  
Qy 5966 TTCAGCGCAGCGCACCAAGTAAAGTGAAGTAACTGCGCTGCTGTTGAAAGGGGCGAGTTT 6025  
Db 1108 CCAGCGATCAGACGCTGCAAGCTGCAAGATCGGCTTCTGCGCGCGGAGCTTGA 1049  
Qy 6026 GCGCATCACTTCGCTGCTGCTCTGCTGAAACCGCGCTCGGCTGCTGTTGCGAGCGTCC 6085  
Db 1048 CGCGCGCATCGCGCTGCGCGCGGCAATCAGCGCGACGCGCTGTGAGCCCGACC 989  
Qy 6086 TCG 6088  
Db 988 GCG 986

RESULT 15  
ARI169831/c  
LOCUS ARI169831 8625 bp DNA linear PAT 17-DEC-2001  
DEFINITION Sequence 1 from patent US 6291204.  
ACCESSION ARI169831  
VERSION ARI169831.1 GI:17907751  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 8625)  
AUTHORS Pasamontes, L. and Tsygankov, Y.  
TITLE Fermentative carotenoid production  
JOURNAL Patent: US 6291204-A 1 18-SEP-2001;  
FEATURES Location/Qualifiers  
source  
1..8625  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 14.8%; Score 1035.4; DB 6; Length 8625;  
Best Local Similarity 56.5%; Pred. No. 3 8e-207;  
Matches 2037; Conservative 0; Mismatches 1536; Indels 30; Gaps 5;

Qy 2486 GTGCCAGCAGCAAGTCGTGCTGGCGAGCGCACTGATGCGCACGCAATACGATGTGATT 2545  
Db 6984 GGGCGCTGCTGAAGGACCGCAAGGGCGGATCGCAATACATGAGCCATGATCTGCTG 6925  
Qy 2546 TTGTTGCTGCTGGAAGTGGCGAATGGCTTGAATGGCTGCTGCTGCTCGCTCAATTCAGCCA 2605  
Db 6924 ATTCGCGGCGCGGGCTGTTCGCTGCTGATGCTGCGGCTTGGCGCTGCGGACCG 6865  
Qy 2606 CAACCTGAATGCTGTTGCTGGAGAGCGATGCGCATCCGCGAGGCAATCATCTGCTG 2665  
Db 6864 GATGCGCGCATCGTATGCTGCGCGGCTCGGCGCTCGGACCGACACCTGGTCC 6805  
Qy 2666 TTTCATCAGCGCATCTCAGCCCGCAACTTCGCTGCTGCTCAACCGCTGATTACCGTG 2725

Db	5804	TGCCACGACACGAGTCTTTCCGCCCGAATGGCTGGCGGGCTGTGCCCCATTGTCGGCGC	6745
Qy	2726	CGTTGGTCAAGTTATCAGGTGGCTTTTCTGCGCTGCGCCGCAATCTGACGGGATAT	2785
Db	6744	GAATGGACGGATCAGAGGTGCGTTTCCGACCAATTCGCGCGCTGACGACAGGCTAT	6685
Qy	2786	TGTTCCATCGCATCAGGCGATTTTGCCCGCATCTTTACGCGCGGATGGTGAAGATCTG	2845
Db	6684	GGCTCGATCGAGGGCGCGCTGATCGGGCTGTGCA-----GGGTTCGATCTG	6634
Qy	2846	TGGAACAACACAGCCGTCACACAGGTAAACACACAGGTGACGCTGGCGGATGGCCGT	2905
Db	6633	CGGTGGATACGATGTCCGAGCTGGAGATACCGGGCGACGCTGACGAGCGGCTCG	6574
Qy	2906	GAATTCGTGCGCAAGTGTGATGATGTCGCGGCTTGACGCCGACGACCATCTGACG	2965
Db	6573	CGGATCGAGGCTGCTGCTGATCGAGCGCGTGGTGGCTCGAGACCCCGCACCTGACC	6514
Qy	2966	CTGGGTTATCAGGTGTTCTTGACACAGAGTGGCAGCTGGCGACGCGCACGCGCTGACG	3025
Db	6513	GTGGGTTTCCAGAAATTCGTGGCGCTGAGATCGAGACCGACGCCCCCATGGCGTCGAG	6454
Qy	3026	CAGCCGATCTGATGATGCCACCGTCGATCAGCAAGCGGTTATCGTTTGTCTACACG	3085
Db	6453	CGCCCGATGATCATGGACGACCGTTCCGAGATGGACGGTACCGCTTCACTATCTG	6394
Qy	3086	CTGCGCTCAGCGCCGATCGGCTATTGATTGAAGATACCCATTAGCTTAACAGCCCGCG	3145
Db	6393	CTGCCCTTCACTCCACCGCATCTGATCGAGGATACGCTACAGCGACGCGCGGAT	6334
Qy	3146	CTGGCGGAGAACACCGCTCGTCAGACATCGCGGACTATGCCAATCAGCAAGGCTGGAGC	3205
Db	6333	CTGAGCATGGGCGCTGGCGCAGGCGTGCCTGGGATATGCCGCCAGCGGGGCTGGACC	6274
Qy	3206	CTGAGTACGCTGCTGCTCAAGACACGCGCATATTACCGATTACCTGAGCGGCAACATC	3265
Db	6273	GGGCGAGGATGGGCGCGAAGGGGCATCTGCCCATCGCGTGGCCCATGACGCCATA	6214
Qy	3266	GATCGATTCTGCAACAGACAGCGCGCCCAAGCTGTCAGCGGCTGCGCGCGGGCTGTTT	3325
Db	6213	GGCTTCTGGCGCGACACACGCGCAGGGGGCGTGCCTGTTGGCTGGGGCGAGGGCTGTC	6154
Qy	3326	CATGCCACACCGGTTACTCTTGGCTCGCGCTGGCGGCTAGCGGAGTTGTAGCAGCG	3385
Db	6153	CACCCCGTCACCGGATATTCTGCTGCCCTATGCGCGCAGGTGCGGATGCCATCGCGCG	6094
Qy	3386	CTGTTGCCACCGATGCCCTCAGCTCAGCCAAACATATCGAACGCTTTCCTCGTCAGCAG	3445
Db	6093	CGCGACCTGACACCGCTGCGCGCTGCGCGGCTGGCGGCTGGGCCATCGATCGCGG	6034
Qy	3446	TGGCGCGAACAGCGATTTTTCGCTGCTGCTAAACCGCATGCTGTTTGGCGGTAAGCGG	3505
Db	6033	GATCGG---ACCGCTTCTGCGGCTGCTGAACCGGATGCTGTCGCGGCTGCCCGCC	5977
Qy	3506	CAGCAGCGCTGGCGGTGATGCAACGTTTTTACCGGCTCGATGCGCGGTTAATTAGCCGC	3565
Db	5976	GACCGTCTGATCGCTGTCAGCGGTTCTACCGCTTCCCGCAGCGCGCTGATCGAGCGC	5917
Qy	3566	TTTTTACCGCGGCACTGCGCTGCGCGATATAACCGCGATTCTGTGCGGCAAGCGCGG	3625
Db	5916	TTCTATGCGCGGCGCTGACATTGGCCGACCGGCTTCGATCGTCACCGGACCGCGGCC	5857
Qy	3626	GTGCCCATCGGTGAAGCGCTGGCGGCTGTTGAATTCGTGGAACACAGGGAAGAAAAA	3685
Db	5856	ATTCCGCTGTCAGCGCCGTGGCTGCTGCCGACGCCCCCTGCTGCGAGGAGAGCA	5797
Qy	3686	TGAAGCGCATTTATGTGATGGGCGAGGCTTTGGCGGCTTGGCGTGGCGATTCGCTGCG	3745
Db	5796	TGAGTTCCGCCCATCGTCACTCGCGCAGGTTTGGCGGGCTTGGCGTTCGCTATCGGCTG	5737
Qy	3746	AGCGCGGGGATACCAACCATCTACTGACGACGCGGACCAAAACCGGCGGACGCGCT	3805
Db	5736	AATCGCGCGCATCGCGACACCATCGTCAGGCGCGGACCAAGCCCCCGCGCGCT	5677
Qy	3806	ATGCTTTGAGGACAGTGGCTTTTACCTTCGATGCGCGGACCCACGGTGATCACCGATCCCA	3865
Db	5676	ATGCTGTGAAACGATCAGGCGCCAGCTCTTCGATGCGAGCCCGACGGTCTGTCAGCCCG	5617
Qy	3866	GGCCCATCGAGAGTTGTTACGCTGGCAGGAAATTCGCTCAGCTCAGCATTAAGCTCGAGCTGA	3925
Db	5616	ACAGCTGGAGAGCTGTGGGCCCTCAGCGGCCAACCGATGGAGCGTGACGTGACGCTGC	5557
Qy	3926	TGCCGGTAAACGCCCTTCTATCGCTGTGTGGGAAAGATGGCAAAACAGCTTGTATTACGACA	3985
Db	5556	TGCCGCTCTCGCCCTTCTACCGCTGACATGCGCGGACATGCGCGGACGCGCGCATGTA	5497
Qy	3986	ATAATCAGCGCTGCTGAGGACGACATCGCCACGTTTAAATCCGCAAGATGTAGAAGCT	4045
Db	5496	ACGACGACGACGAGCTGATCCGCCAGTCCCTTCAATCCCGCGGATGTGATGGCT	5437
Qy	4046	ATCGTCAATTTCTTCCTTATTCACCTGTAAGTATTTAGAGAGGTTATCTGAAACTCGGCA	4105
Db	5436	ATTCGCGCTTCCACGATTACGCGGAGGTTTATCGGAGGGGTATCTGAAAGCTGGGA	5377
Qy	4106	CGGTCCGCTTCTGCAAGTGTGACATGCTGCGGCTGCGCGCGCAGTTGGGACGCTGTC	4165
Db	5376	CCACGCCCTTCTGAAGCTGGCCAGATGCTGAACGCCGCGCGCGCTGATCGGCTGTC	5317
Qy	4166	AAGCATGGCGCAGCTCTACAGATGGTGGGAAATTTATTAGGACGATCATCTGCGTC	4225
Db	5316	AGGCATACCGCTCGGTCACAGCATGGTGGCGCTTCAATCAGGACCCGCTCTGCGGC	5257
Qy	4226	AGCGCTTTCTTCCACTCATTTGCTGGGCGGTAACTCTTTGCAAGCTCATCGATCT	4285
Db	5256	AGGCCCTTCTCGTTCCACACGCTGCTGGTGGGCGGAAACCCGTTTCGACAGCTCGATCT	5197
Qy	4286	ATACCTTAATTCATCGCTGAGCGTGAATGGGGCGTGTGTTTCGCGCGCGGCAACCG	4345
Db	5196	ATGCGCTGATCATGCGCTGGAAACGCGCGCGCGCTGTGGTTTCGCAAGGCGGACCA	5137
Qy	4346	GGCGCTGCTGAGGCGATGGCGCACTGTTTCAGGACTTGGGGCGCGAGCTGTGTACTGA	4405
Db	5136	ACCAGCTGCTCGCGGCTGCTGCGCTGTTTCAGCGCTTTCGCGCGCACGCTGCTCTGA	5077
Qy	4406	ATGCGGAGTGAAGCCAGCTGGAAACAGCGCAATCGCATTAGCGCGCTTCAGTTAGAGG	4465
Db	5076	ATGCGCGCTGACGCGGATCGACACCGAGGGCGATCGCGCCACGGCGCTCACGCTCTGG	5017
Qy	4466	GGGACGACGCTTCGATGCGCGCTGTGGCTTCAATGCCGACGCTGTGCTATACCTAGC	4525
Db	5016	ACGGGGCGAGTTGCGCGGATACGTTGGCCAGCAACGGCGACGTTGTCACAGCTATC	4957
Qy	4526	ACAACTGCTTGCACCATCGCTGGCAATGAAGCTGCGACATCGCTGAAGCGTAAAGC	4585
Db	4956	GGGACCTGCTGGGCCATACCCGCGCGGCGCACCAAGCGCGGATCTCTGAACCGGACG	4897
Qy	4586	GCATGAGCAACTGCTGTTGTTGCTATTTTGGCTGATCAGCGCGCTGATGACAGCTCG	4645
Db	4896	GCTGTGATGTCGCTGTCGCTGTCATTTTCGCGCTGTCGCGGCGCCCGGAGAACTGG	4837
Qy	4646	CCACACACACGCTGTTTGGCCGCGCTTATCGTGAAGTTGATCGATGAGATTTTCAACA	4705
Db	4836	CCACACACAGCTGATCTTCGCGCGGCTACAGGGGCTGTGAACGAGATCTTCACAG	4777
Qy	4706	GCAGCAGCTGGCAGACGATTTTTCATTTTACCTGCAACGCGCTTGCAGACGATCCGT	4765
Db	4776	GGCCACGCTGCGGACGATTTTCGATGATCTGATTCGCTTTCGCGCTGCGTACCGATCCCA	4717
Qy	4766	CGCTGGCACCGCGCTGCGGACGCTTTTATGTTAGCGCGGCTGCGGATCTCGGCA	4825
Db	4716	GCCTGGCCCCGAGGGGATGTCCACGCTTACGCTCTTGGCGCGCTTCCGCTCTGGGCC	4657
Qy	4826	CGCTGACATCGACTGGCAACAGGAAGGACCGCGCTTGGCGGATCGAATTTTGTCTATC	4885
Db	4656	GGCGCGATGTCGATTGGGAGCGGAGGCCCGGCTATCGCGAGCGCATCTTCGAGGAAC	4597

4886	Qy	TGAGCAGCACTACATCCCGGGATTACGTGAGCAATTTAGTGCACACAGAAATGTTTACGC	4945
4596	Db	TGGAGCCCGCGCCATCCCGACCTGCGCGAAGCACCTTGACCGTACGCGCGATCTTCAGCC	4537
4946	Qy	CGTTTGATTTTCGGACACGCTGATCCCATCACGGCTCGGGTCTTTTCGTGCGAGCCGA	5005
4536	Db	CCGCCGATTTACGACCGAACTGTCGGCCCATACGGCAGCGCTTCTCGTTCGAGCCGA	4477
5006	Qy	TTTTGACGCAAGCGCTGTTTCGCGCCGATAAACCGCGATGCCGATATCAGCAATCTCT	5065
4476	Db	TCCTGACGCAATCCGCTGTTTCGCGCGCATAAACCGCAGCCGCGCATCCCGAACTTCT	4417
5066	Qy	ATCTGTGGGTTCGGTATACGATCCAGCGCGGGCGTCCCGCGTGATCGGTTTCGGCCA	5125
4416	Db	ACATCTGGGGGCGGCGACGATCCGGTGGGGCATCCCGGGTGTCTGTCGACGCGCCA	4357
5126	Qy	AGGCCACCGCAGGCTGATCTGAGGATCGCGCGGAATGAATCGACAGGCTTTTACTTGA	5185
4356	Db	AGGCCACGGCGAGTCTATGCTGCGACCTGGCCGTCGATGACCGATCTGACGGCGAC	4297
5186	Qy	GCAAGTAACGAAACCATGCGGGTGGCTCGAAGAGTTTTCGCCACCGCCGCGAAGCTGTT	5245
4296	Db	TTCCGAAGCG---GCCATCGCGCAGGGTTCGCAAGCTTCGCGCAGCGCGCCAAAGCTGAT	4240
5246	Qy	TGATGACCCAGCGCGCGCAGCAGCTGATGCTGATGCGTGGTGTGCTCACTGCCGATGA	5305
4239	Db	CGCGCCCGCATCCGCGAGGATACGCTATGCTCTATGCTGTGTGAGGATATGGGATGA	4180
5306	Qy	TGTGATTGATGGGCAAAAGCTTGGCGGAAGCGGCACGACGATGCCGTGCAAGACGCGCA	5365
4179	Db	CGTGATCGACGGGCAGGTGATGGGTTCTGCCCCCGAG-----GCGGGCGCGACCCACA	4126
5366	Qy	GGCAGCTATCAGCATCTGCATAATTGAAACCCGCGCGCTACAGCGCGCGCACATGGA	5425
4125	Db	GGCGCGCTGGGGCGCTGCGCGCGACACGCTGGCCGCGTGCACGAGGACGGCCCGAT	4066
5426	Qy	TGAACCGCGCTTTAGGCGGTTTCAGGAAGTGGCGATCATTCACAGTGCGCGCAACAAT	5485
4065	Db	GTCGCCCTTCGCGCGCTGCGCCAGGTGCGCCGCGCGCATGATTTCGCGACCTTTG	4006
5486	Qy	GGCGTTTTCATCTGGAAGCTTCGCTATGGATGCACGCAACGAAATTAACGCGAGCTT	5545
4005	Db	GCCGATGGACCTGATCAGAGGTTTCGGGATGGATGTTCGGGATCGCGAATACCGACGCT	3946
5546	Qy	CGATGACACGCTGTGTTACTGCTATCACGTGCGGGCGTGGTTCGGTTTGATGATGGCGG	5605
3945	Db	GGATGACGTGCTGAATATTCCTACCAAGTCGCGGGGTGTCGGCGTGTGATGATGGCGG	3886
5606	Qy	CGTAATGGGCTGGCGACGAGAGGGTGTCTCGATCACGCTGCGATTTAGGACTGGCGTT	5665
3885	Db	GGTGATGGGCTGCAGACGATGCGGTGCTGGATTCGCGCTTCGCGATCTGGGCTTCGCTT	3826
5666	Qy	CCAGCTCACTAACTTGGCGCGACATTGTAGAAAGATGCCGAAATGGTTCGTCTATCT	5725
3825	Db	CCAGCTGACGNAATCATCGCTCGGACGTGATCGACGATCGCGCATCGGCGCTGATCT	3766
5726	Qy	GCCGCAATCTCTGGCTCGATCAGCGGGATATACGCGCGATACGCTGATGCACGCAACA	5785
3765	Db	GCCTGCCGCACTGGCTGGCGCAGGGGG-----GGCGACGGTTCAGGGTTCGGTGCC	3715
5786	Qy	TCGTGCGACGCTCGCTCACTGGCAGCGCTTAGTGGCGGAGCGCAACCTATTATCA	5845
3714	Db	TTTCGAGCGCTCTATTTCGTCATCCGCTGCTTGAACGGCGCAGCCCTATTATGC	3655
5846	Qy	CTCGGCGCGATTCGGGTTTACCGGGTTTACCGCTGCGCTCGGCGTGGGCCATTCGCTACGC	5905
3654	Db	CTCGGCGCGCAGGGCTTCGCGATCTGCCCGCGCTGCGCGTGTGATCGCGCGCGC	3595
5906	Qy	TCGCGGGCTTTATCGGAAATTGGCGTCAAGTTTCAGCACGCCGGTGTGCACCGCTGGGA	5965
3594	Db	GCTCGCTATCTATTCGCGCAATTCGGGACGCGCATCCGACGGTGGCCCCGAGGCTATATCG	3535
5966	Qy	TTCAAGCAGCGCACCAAGTGAAGGTGAAAACTGGCGCTGCTGCTGTGAAGAGGGCAGGTTT	6025

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 24, 2005, 22:47:11 ; Search time 2935 Seconds  
(without alignments)  
12518.125 Million cell updates/sec

Title: US-10-808-979-18  
Perfect score: 6999  
Sequence: 1 gcagatggcaatgggaatgt.....gacgttttagacatgcaata 6999

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq 23Sep04:.\*  
1: Geneseqn1980s:.\*  
2: Geneseqn1990s:.\*  
3: Geneseqn2000s:.\*  
4: Geneseqn2001as:.\*  
5: Geneseqn2001bs:.\*  
6: Geneseqn2002as:.\*  
7: Geneseqn2002bs:.\*  
8: Geneseqn2003as:.\*  
9: Geneseqn2003bs:.\*  
10: Geneseqn2003cs:.\*  
11: Geneseqn2003ds:.\*  
12: Geneseqn2004s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2623	37.5	6918	2	AAQ06299 Sequence
2	2619.6	37.4	8609	12	Adq14666 Reporter
3	2619.6	37.4	8609	12	Adq26394 Carotenoi
4	2619.6	37.4	8609	12	Adq77262 Nucleotid
5	2619.6	37.4	8609	12	Adq48673 Plasmid p
6	2135.2	30.5	12753	12	Adm98599 Geranylge
7	1900.6	27.2	4624	12	Adp74122 Pantoea a
8	1618.8	23.1	7494	8	Abz69177 Vector co
9	1618.8	23.1	8547	8	Abz69178 Vector co
10	1374.8	19.6	110000	10	Continuation (42 o
11	1374.8	19.6	110000	10	Continuation (7 of
12	1374.8	19.6	110000	10	Continuation (8 of
13	1037	14.8	8625	2	Aav40146 Flavobact
14	1037	14.8	11233	2	Aav40151 DNA seque
15	1035.4	14.8	8625	2	Aat45143 Flavobact
16	827.4	11.8	1493	2	Aax90713 Erwinia u
17	826.2	11.8	1479	2	AAQ06296 Sequence
18	826.2	11.8	1479	2	Aav84081 Carotenoi
19	826.2	11.8	1479	2	Aax19119 Erwinia u
20	826.2	11.8	1479	12	Ado05014 P. anan
21	808.6	11.6	1479	6	Aad35512 Pantoea s

ALIGNMENTS

RESULT 1  
AAQ06299  
ID AAQ06299 standard; DNA; 6918 BP.  
XX  
AC AAQ06299;  
XX  
DT 24-OCT-2003 (revised)  
DT 28-JAN-1991 (first entry)  
XX  
DE Sequence encoding six enzymes of the carotenoid biosynthetic pathway.  
XX  
KW Carotenoid biosynthesis; vitamin A; cancer; food coloring; ss.  
XX  
OS Pantoea ananatis.  
XX  
FH Key Location/Qualifiers  
FT mat\_peptide 225..1130  
FT /tag= a  
FT /note= "See AAQ06293"  
FT mat\_peptide 1142..2435  
FT /tag= b  
FT /note= "See AAQ06294"  
FT mat\_peptide 2421..3567  
FT /tag= c  
FT /note= "See AAQ06295"  
FT mat\_peptide 3582..5057  
FT /tag= d  
FT /note= "See AAQ06296"  
FT mat\_peptide 5096..5983  
FT /tag= e  
FT /note= "See AAQ06297"  
FT mat\_peptide complement(5928..6452)  
FT /tag= f  
FT /note= "See AAQ06298"

22 808.6 11.6 1479 8 ABT14193  
23 808.6 11.6 1479 8 ACC44762  
24 808.6 11.6 1479 12 ADQ14630  
25 808.6 11.6 1479 12 ADQ77205  
26 808.6 11.6 1479 12 ADQ48640  
27 713.4 10.2 1518 2 AAQ13719  
28 713.4 10.2 1518 2 AAT40792  
29 713.4 10.2 1518 2 AAT37094  
30 713.4 10.2 1518 2 AAT91544  
31 708.6 10.1 1482 10 ACF70993  
32 708.4 10.1 1517 2 AAQ13720  
33 708.4 10.1 1522 2 AAT40793  
34 708.4 10.1 1522 2 AAT37095  
35 708.4 10.1 1522 2 AAT91546  
36 692.6 9.9 1521 10 ADC99016  
37 622.8 8.9 1497 2 AAV73181  
38 515.6 7.4 1198 2 AAQ13718  
39 515.6 7.4 1198 2 AAT40791  
40 515.6 7.4 1198 2 AAT37093  
41 515.6 7.4 1198 2 AAT41743  
42 515.6 7.4 1198 2 AAT91543  
43 461.4 6.6 1581 2 AAT31798  
44 449.6 6.4 2034 11 ADQ96835  
45 449.6 6.4 2051 11 ADQ96934

ABT14193 Pantoea s  
ACC44762 Pantoea s  
ADQ14630 Pantoea s  
ADQ77205 Nucleotid  
ADQ48640 Pantoea s  
AAQ13719 Phytoene-  
AAT40792 Phytoene-  
AAT37094 Phytoene  
AAT91544 Erwinia h  
ACF70993 Photorhab  
AAQ13720 Phytoene-  
AAT40793 Phytoene-  
AAT37095 Phytoene  
AAT91546 Erwinia h  
AAV73181 C. utilis  
AAQ13718 Phytoene  
AAT40791 Phytoene  
AAT37093 E. herbic  
AAT41743 Phytoene  
AAT91543 Erwinia h  
AAT31798 Erythroba  
ADQ96835 Crticrt2  
ADQ96934 Crticrt2

Wed Jan 26 08:20:42 2005

PA (KIRI ) KIRIN BEER KK.  
 XX Misawa N, Kobayashi K, Nakamura K;  
 XX WPI; 1990-322212/43.  
 DR DNA sequences encoding enzymes for carotenoid biosynthesis - for prodn.  
 XX of carotenoid cpds. e.g. beta-carotene, lycopene, phytoene, etc.  
 PT Disclosure; Fig 7; 40pp; English.  
 XX  
 CC Gene products are useful for the synthesis of carotenoids, useful as food  
 CC coloring, vitamin A precursor, and possibly in prevention of cancer. See  
 CC also AAQ06293-9. (Updated on 24-OCT-2003 to standardise OS field)  
 XX  
 SQ Sequence 6918 BP; 1522 A; 1876 C; 1881 G; 1639 T; 0 U; 0 Other;  
 Query Match 37.5%; Score 2623; DB 2; Length 6918;  
 Best Local Similarity 64.6%; Pred. No. 0;  
 Matches 4006; Conservative 0; Mismatches 2170; Indels 21; Gaps 6;  
 400 CTTGAACATTACTGCTGCCGACAGCAAAAGCGATGCGTGGCGCGGATGCGTGC 459  
 DB 303 CTTGATCAGTTATTGCCCGTGAGGAGAAACGGGATGTTGGGTGCCGCGATGCGTGA 362  
 460 GGAAACGCTGGCAGGCGAAAGATTCGTCCTTTTACTCTGCTGCGACGCGCGAT 519  
 DB 363 GGTGCGCTGGCAGCCGCGGAAACGATTCGCCCCCATGTTGCTGTGTGACCGCCGCGAT 422  
 520 ATGGGTGGCGAGTGACGCAAAATGGCGTTCTCGATCTCGCTCGCTGCGAGTGGAAATGTTG 579  
 DB 423 CTGGGTGGCTGTAGCCATGAGCGATTACTGGATTTGGCTGTGCGGTGGAATGCTC 482  
 580 CACGCGGATTCGCTGATTCGATGATCATTCCTCGATGATGATGATGATGATGATGATGATG 639  
 DB 483 CACGCGCTTCGCTGATCTTGCAGATATGCCCTGATGATGATGATGATGATGATGATGATG 542  
 640 GGTGCGCTTACGCTGATCGCAATTTGTTGAAACGTTGCGGATTTGCGCTGCGGATTCGCG 699  
 DB 543 GGAAGCTTACCATTTCTTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 602  
 700 CTCTTTAGCGCGCATTTGAAGTATGTCATTCGACCCGTTGCTGCTGCTGCTGCTGCTGCTG 759  
 DB 603 TTGCTGATGAAGCTTTGGCGTAAATGTCGATGATGATGATGATGATGATGATGATGATG 662  
 760 TCTGAAGCATTCGTGATCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 819  
 DB 663 AATCGGCGGTTTCTGAACCTGTCAACGCGCATCGGATGATGATGATGATGATGATGATG 722  
 820 TTCCAGGATCTGACACGCGCAGCAGCGCGCAGCGCGCGGAGCGATCGCCATGACCAAC 879  
 DB 723 TTTCAAGGATCTGTGAAGGGATAGCGCGCAGCGCTGAAGCTATTTTGTATGACGAAT 782  
 880 GAACTGAAACCGAGCTGTGTTTGGCGCAGCGTGCATGCGTGAATGCGGCGATGCGCTGAC 939  
 DB 783 CACTTTAAACCAAGCAGCTGTTTGTGCTTCAATGATGATGATGATGATGATGATGATGATG 842  
 940 GCTTCAAGGATGCGGCAAGACTTTAGCTTCTTCCCGCAGGATTTGGGCGCAGGCGTTT 999  
 DB 843 GCCTTCAAGGATGCGGCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 902  
 1000 CAATGCTGCAACGCTCGCGAGCGTTGCAAAACACACCGGTAAGATGTCACAGGAT 1059  
 DB 903 CAATGCTGCAACGCTCGCGAGCGTTGCAAAACACACCGGTAAGATGTCACAGGAT 962  
 1060 CAGGCAATTCACGCTGTTACAGATGCTCGGTGCTGACGCGCGGAAACGTCGCTGCGC 1119  
 DB 963 GCCGGTAAATCGACGCTGGTCAATCTGTTAGGCGCGAGCGGCTGTTGAAGACGCTGAGA 1022  
 1120 GATCACTGCGCAGCGGATGACACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1179  
 DB 1023 CAACATCTTACGCTTGCAGTGAGCATCTCTCTGCGGCTTGCACCAACGCGGCGACGCGCACT 1082

QY 1180 CGCAATATATGACGCGCTGTTTAAATCAACAGCTAGCGATATTAACATGAGCGCGCTC 1239  
 DB 1083 CAACATTTTATTGAGGCTGTTTGAACAAAACCTGCTGCGCTAGTTAAAGATGCTGC 1142  
 QY 1240 AGCGGTGGGCACTTTGGGTGATCGCGCGCGCTCTACAGCCACTTTACGCGCTGC 1299  
 DB 1143 A-----TGAGCCATTTGCGCGGATCGCACCGCTTTTACAGCCATGTTTGGCGATTAC 1197  
 QY 1300 AGCGGTAGCACAAACGCTGCTGCGCGCGGCGATCGCATCATCATTCATCAGCAAGCGC 1359  
 DB 1198 AGAATCTCGCTCAGAACTGCTGCGCGCGGTCACTGGGTGACCTTTATTCAGCAATACG 1257  
 QY 1360 ATGCCCGCACTTTGCTTAGCGACGAACGATCGAATTTGTTGCTGCGCGCAACAGCGC 1419  
 DB 1258 ATATTAACACTTGTATCGATAGCAAAACCATTTGATTTCACTCGTGGGACAGACGCC 1317  
 QY 1420 ATCTGCGGTTGCTGCGCGCGCTGTTGATCGGCTGCGCTGCGCGCGCTGCTGCGC 1479  
 DB 1318 ATCCCGCGCGCTTAAACGCGCGTGTACACCTGCGCGCTCATCTCTGCGCGCTCA 1377  
 QY 1480 TGTTCGCGTATCGACGATCTCGCGTCTGCGACCGATATGCTGTGCGCGAACTCCGCC 1539  
 DB 1378 TGCTGAAGCTCATCAATGAATGCGCGCACACCGGATATGCTGTGCGCGAACTCCGCC 1437  
 QY 1540 CGGTACTGAAGCAATTGAACATCGATGCGGTGATGCGCGCAAGATGGAAGCGCGCGC 1599  
 DB 1438 AGGCATTTAAGCATCTGCGCGTGTGATGCGGTGATTTGTTGATCAAAATGGAACCGCGCGC 1497  
 QY 1600 GATTTGCTGCTGAAGCGCTGATCTGCGTTTGTTCGCTGCGCTGCGCTGCGCTGCTGCA 1659  
 DB 1498 CGCTGCTGCTGGAAGCACTGGGACTGCGTTTATCTCTGCTGCGCTGCGCTGCTCTCA 1557  
 QY 1660 ATCGTGAAGCGCGGATTCGCTTCCGCTGCGGTGATGCGCTTCCGTTTGTGCAAGATGACAAG 1719  
 DB 1558 ATCGTGAACCGGATGATGCGCTGCGGTGATGCGCTTTCGAATACGGGACCGACCGCG 1617  
 QY 1720 CGCTGAACGTTTTCGCGCAGCAGCATATCTATGATCGCATGCTGCTGCTGCTGCGCG 1779  
 DB 1618 CTGCGAAGCTTATGCGCGCAGTGAAGAAATTTATGATGCTGCTGCTGCTGCTGCTGCTG 1677  
 QY 1780 ACGTGATCTCAAAACACGCGCGGCTTTAATTTGACGAGCGCGCGGATTAATCATAGT 1839  
 DB 1678 GTGTCATTTGCCGACACACAGCCACAGATGCGCTTAGCCCCCGCAAAAGCTTCAACAGT 1737  
 QY 1840 GCTGTCGCGCTGCGCAAAATCAGCAGATGCTGCGCGCTTGTGATTTTCCAGCTGACG 1899  
 DB 1738 GTTTTTCGCCACTGCGCGCAATCAGCCAGCTGTTTCTGAACTGGAATTTTCCCGCAAG 1797  
 QY 1900 AACTGCCGCTGCTATCACCGCGTGGGCGCATCTCCGCGCGCGCTTCTCTGCGCGC 1959  
 DB 1798 CGTTTACCGCTGTTTTCATGCGCTGCGGCTCTGCGCGGAAACGCGCACCGCTCAACGT 1857  
 QY 1960 TCCATGCGCCCTG---GCCAGCGCTGCTGACGCGGTGTTTATGCTGCTGCTGCTGCTG 2016  
 DB 1858 CTTTCATCCGTTTATTTTATCTCCTCAGAAACCCCGGATTTTCCGCTGCTGCGCGC 1917  
 QY 2017 TCAAGGCGATCGCTTCCGCTGCTTCTGCTGCGCGCGCTGCGCGCGCTGCGCGCTGCGCG 2076  
 DB 1918 TTCAGGCGACCGCTTATGCGCTGTTTAAACGATAGTGAAGCTTGTGAAGAAATGAGC 1977  
 QY 2077 TATCGCTGCTGATCGCCCATTTGTGGGGATTTAAACCGCGCAACAGACGATCATGCTGAGC 2136  
 DB 1978 GTGAGCTCTGTTAGCCCATCTGCTGCTTTCAGGACTCTCAGTGTGAAGAGCTGGCGC 2037  
 QY 2137 TCGTGGCGCGGCTGGGTGACGATTTGCTGATGACGCGCGCGCTTACAGCAGCGC 2196  
 DB 2038 GAAGCGCTCATACAGAGTGGGATTTTTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2097  
 QY 2197 AGCTGTTTATCTCATGCTGCGGTTTAAACAGCGCGCTGGAAGCACTGGAATCGGTACGC 2256  
 DB 2098 AGCTGCGATTCACCGCGCGGATGATACGCTGCTGCGCGGATTAATTCAGGACGC 2157  
 QY 2257 CGATGCTGCGCTGCGGATTTGCTTTGATCAGCGCGCGCTGCGCGCGCTGCTGAGTGGC 2316





QY 4477 TTCGATCGCGCGTGTGGCTTCCAAATCGCGAGTGGTGATACCTACGACAAACTGCTT 4536  
DB |||||  
QY 4374 TTCGAGCGAGCGGTGCGGTCAAAATGAGATGGTGTACATCTATCGGAGCTGTGA 4433  
DB |||||  
QY 4537 GCGCACCAATCGGCTGGCAATGAAACGTGCGACATCGCTGAAGCGTAAAGCGATGAGCAAC 4596  
DB |||||  
QY 4434 AGCAGACCCCTGCGCGGTAAAGCAGTCCAAACAACTGCAGACTAAAGCGATGAGTAA 4493  
DB |||||  
QY 4597 TGGCTGTTGATCTTATTTTGGCTGAAATGAGCGGATGAAACAGCTCGCGACCAAC 4656  
DB |||||  
QY 4494 TCTGCTGTTGCTCTATTTTGGTGTGAATCAACCATCATGATCAGCTCGCGCATCACAG 4553  
DB |||||  
QY 4657 GTCTGTTTGGCGCGGTATCGTCAAGTTGATCGATGAGATTTTCAACAGCAGCCAGCTG 4716  
DB |||||  
QY 4554 GTTGTGTTGGCGCGCGGTAAACGCGAGCTGATGAGAAATTTTAAATCATGATGGCTC 4613  
DB |||||  
QY 4717 GCGACGATTTTTTCACTTTACCTGACGCGCCCTCGACAGCGATTCCTGCTGCGACCG 4776  
DB |||||  
QY 4614 GCGAGGACTTCTCACTTTATCTGACGCGCCCTGTGTACGGAATTCGTCACTGGCGCT 4673  
DB |||||  
QY 4777 CCGGCTGGCGAGCTTTTATGTTTGGCTGAGCGCGGTGCGCATCTCGGACCGCTGACATC 4836  
DB |||||  
QY 4674 GAAGTGTGGCGAGTACTATGTGTGGCGCGGTGCGCATTTAGGACCGCGAACCTC 4733  
DB |||||  
QY 4837 GACTGCAACAGAAAGGACCGGCTTGGCGATCGAAATTTTGTCTTATCTTGAGCAGCAC 4896  
DB |||||  
QY 4734 GACTGACGTTGAGGGGCAAACTACGCGACCGTATTTTGGCTGCTTGGAGCAGCAT 4793  
DB |||||  
QY 4897 TACATCGCGGATTAAGTCAGCAATTAAGTGAACACAGAAATTTTACGCGGTTGATTT 4956  
DB |||||  
QY 4794 TACATGCTGGCTTACGAGTACGAGTGTGTGAGCGGATTTTACGCGGTTTGAATTT 4853  
DB |||||  
QY 4957 GCGGACACGCTGATGCCATCAAGCTCGGCTGCTTGGCTGAGCGGATTTTGAACGAA 5016  
DB |||||  
QY 4854 GCGGACCACTTAATGCTTATCATGGCTCAGCTTCTTGTGAGCGCGGTTCTTACCCAG 4913  
DB |||||  
QY 5017 AGCGCTGTTGCGCGCGATTAACCGCGATGCGATATCAGCAATCTCTATCTGTGGGT 5076  
DB |||||  
QY 4914 AGCGCTGTTTGGCGCGATTAACCGCGATTAACCAATTAATCTTACCTGTGGCTG 4973  
DB |||||  
QY 5077 GCGGTAAGCTGAGCGCGGCGGTGCGCGGATCGGTGATCGGTGCGGCAAGCGCC 5136  
DB |||||  
QY 4974 GCGGACGATCCCGCGCGAGCATTCCTGCGCTCATCGGCTCGGCAAAAGCGACGA 5033  
DB |||||  
QY 5137 AGCTGATGCTGAGGATCGCGCGAATGAATCGACAGCTTACTTGAACAGTAAACGC 5196  
DB |||||  
QY 5034 GGTGATGCTGAGATC---TGATTTGATTAATCCGTGTTACTCAATCATGCGGTG 5090  
DB |||||  
QY 5197 AAACCATGGCGGTGGCTCGAAGATTTCCGCAACCGCGCAAGCTGTTTGAACCGA 5256  
DB |||||  
QY 5091 AAACGATGCGAGTTGGCTCGAAAGTTTGGAGACGCTCAAAGTTATTTGATGCAAAA 5150  
DB |||||  
QY 5257 CCGCGCGAGCAGCTGATGCTGTATGCGTGGTGTGCTCACTCGGATGATGATGATG 5316  
DB |||||  
QY 5151 CCGCGCGAGCTGATGATGCTCTACGCTGCTGCGCGCATTTGTGACGATGTTTGAAG 5210  
DB |||||  
QY 5317 GCGAAACGCTGGCGAGGCGGACGAGCATCGCGTCAAGACGCGGAGGACGATGCG 5376  
DB |||||  
QY 5211 ATCAGAGCTGGGCTTTCAGGCCGCGAGCGCTGCTTAACAAGCCCGAACAAGCTG 5270  
DB |||||  
QY 5377 AGCATGCAAAATGAAACCCGCGCTTACAGCGCGCGCATGATGATGAAACCGGCT 5436  
DB |||||  
QY 5271 TGCAACTTGATGTAAGAAACGCGCGAGCGCTTATGAGGATCGCATGCAACCGCGCT 5330  
DB |||||  
QY 5437 TTAGGCGTTTACAGGAGTGGGCTCATTTACAGCTGCGGCAACAACTGGCGTTTATG 5496  
DB |||||  
QY 5331 TTGCGGCTTTTCAGGAGTGGCTGATGCTCATGATATGCGCCCGGCTTACCGCTTATG 5390  
DB |||||  
QY 5497 ATCTGAAGGCTTTCGCTTATGATGACGCAACGAACTTACGCGAGCTTTCGATGACACGC 5556  
DB |||||  
QY 5391 ATCTGGAAGGCTTCGCGATGATGATACGCGAAGCGCAATACAGCCAACTGGATGATACGC 5450  
DB |||||

QY 5557 TGGCTTACTGCTATCACTGCTCGCGGCGTGTGCTGGTGTGATGATGCGCGCGGTATGCGCG 5616  
DB |||||  
QY 5451 TGGCTTATGCTATCACTGCTGCGAGCGTGTGCTGGCTTGAATGCGCAAAATATGCGCG 5510  
DB |||||  
QY 5617 TGGCGAGCAAGCGGTGCTGATCAACGCTGCGGATTTAGGACTGCGGCTTCAGCTCACTA 5676  
DB |||||  
QY 5511 TGGCGGATAACGCGCACGCTGGAACGCGCTGTGACCTTGGGCTGGCATTTGATGACCA 5570  
DB |||||  
QY 5677 ACATTTGCGCGCATTTGTAGAGATCCGAAATGCTGCTGCTATCTGCGCGAAATCCT 5736  
DB |||||  
QY 5571 ATATTTGCTCGGATTTTGTGACGATGCGCATGCGGCGCTGTGTTATCTGCGGCAAGCT 5630  
DB |||||  
QY 5737 GGTCTGATCAGCGCGGATTAACGCGCGGATACGCTGATGTCACACCGCAACATGCTGCGAGCG 5796  
DB |||||  
QY 5631 GGTCTGAGCATGAAAGTCTGAAACAAAGAAATTTATGCGGCACCTGAAACCCGTCAGGCGC 5690  
DB |||||  
QY 5797 TCGCTCACTGCGAGCGGTTTACTGCGGAGGCGGAACCTTATATCACTTCGCGCGGAT 5856  
DB |||||  
QY 5691 TGAGCGGTATCGCCCGTGTGTTGTTGAGAGAGCAGAACTTACTATTTGCTGCGCACAG 5750  
DB |||||  
QY 5857 CCGGTTTACCGGTTTACCGCTGCGCTCGGCGTGGGCCATCGCTACGCTCGCGCGGCTT 5916  
DB |||||  
QY 5751 CCGGCTTGGCAGGTTGCCCCCTGCTGCTGCGCTGGGCAATCGCTACGCGAAGCAGGTTT 5810  
DB |||||  
QY 5917 ATCGGAAATTTGGCGTCAAAAGTTTCAACGCGCGGTGTCAGCGCTTGGGATTCAGGCGAGC 5976  
DB |||||  
QY 5811 ACCGAAATAGGTTTCAAAAGTTTCAACAGGCGCGCTCAGCAAGCGCTGGGATCAGCGGCGAGT 5870  
DB |||||  
QY 5977 GCACAGTAAAGTTGAAACCTGCGCTGCTGCTGTAAGGGGCGAGGTTTGGCGATCACTT 6036  
DB |||||  
QY 5871 CAACGACACCGCGGAAATTAACGCTGCTGCTGGCGCGCTTGGTCAGGCGCTTACTT 5930  
DB |||||  
QY 6037 CCGCTGCTGCTGCTGCTGAAACCGGCTCGGCTGCTGCTGTCAGCGCTTCTCGTTGATTTT 6096  
DB |||||  
QY 5931 CCGGATGCGGCTCATCTCCCGCTGCGCATCTCTGGCAGCGCCGCT---TCT 5984  
DB |||||  
QY 6097 AGCTCGTGAAGCTGCGCGAGCGTGGCTTGCAGCTTATTCAGCGGTGCGCGTGAAGAA 6156  
DB |||||  
QY 5985 AGCGCATGCTTTCCCGGAGCGTGGCTGAAGTTTGAACAGGGGCGCGCATAGAGAA 6044  
DB |||||  
QY 6157 ACCAAACGACAGCGACCTTCAACGCGCGCACCGCATGATGATGCGGTGCGCATGTA 6216  
DB |||||  
QY 6045 GCGAAAGAAACACACCTTCTTTGCGCTTACGCGGTGATGATGATGCGCATATA 6104  
DB |||||  
QY 6217 TAAGCGCTTAAGATAGCTTTGCGCGGATATAGGGAAACCGCGAGCTTGAATGACACAG 6276  
DB |||||  
QY 6105 CAACCGTTTGAAGTAGCGCTTTCGCTGGAATATAGCGGAATGCGCAACGTTGATGACACAG 6164  
DB |||||  
QY 6277 GCGATGCGACCATGAAGTAGAGCGCGCTGCTGCTGCTTCCGCGCAACCAATCCACTG 6336  
DB |||||  
QY 6165 CCGCTGCTGACCATTAATTAAGATGATTCATACGCGCTTACCTGCGCATTCCTG 6224  
DB |||||  
QY 6337 CAGCGCGCACATGCTTGCACACGACATAAATACGCAATTCGCGAGTACCGCAAAACAC 6396  
DB |||||  
QY 6225 GAGCGCGCACATTCCTGCTACTGCCAGATAAATCAGAGGATCGATAATGACGCAAAAC 6284  
DB |||||  
QY 6397 CACCGCATTAAGATGCTGTTGAGCTTAACTTACCGCTGTCGCGTTTCAATGCTGCGAGATG 6456  
DB |||||  
QY 6285 CACGCGATAAAGATCGTTAACTTCAACGCGACCTTTTACCGGTTTATGATGTAAGATG 6344  
DB |||||  
QY 6457 CCAGCCCATCCCAACCGTGCATGATGTTTATGCGACAGCGCGCTTACGATTTCCAT 6516  
DB |||||  
QY 6345 CATCCCAACCCAGCGCTGATGATGTTTGTGCGCAGTGCAGCAATCACTTCCAT 6404  
DB |||||  
QY 6517 CACCAACCGTTGCGCAACAGATAAGCAAGCTTCCATTAACAGAGATTTGCTGCTCAT 6576  
DB |||||  
QY 6405 GCGAATCAGGTAACGAAACCGATCAGGCGCATTCGAAATCCCAACATAATTTCTCCGCT 6464  
DB |||||  
QY 6577 TGTGAAAGGGAAGTA 6593  
DB |||||  
QY 6465 AGAGAGCTTGGCAGCA 6481  
DB |||||

## RESULT 2.

ADQ14666  
ID ADQ14666 standard; DNA; 8609 BP.

AC AC  
ADQ14666;

XX XX  
23-SEP-2004 (first entry)

XX XX  
Reporter plasmid pPCB15 nucleotide sequence SEQ ID NO:43.

XX XX  
carotenoid; carotenoid overproducing microorganism;  
KW functional isoprenoid enzymatic biosynthetic pathway; plasmid; circular;  
KW gene; ds.

XX OS  
Synthetic.

XX XX  
WO2004056974-A2.

XX PD  
08-JUL-2004.

XX PF  
19-DEC-2003; 2003WO-US041811.

XX PR  
19-DEC-2002; 2002US-0435612P.

XX XX  
(DUPO ) DU PONT DE NEMOURS & CO E I.

XX XX  
Cheng Q, Rouviere PE, Tao L;

XX PI  
WPI; 2004-525436/50.

XX XX  
Novel carotenoid overproducing microorganism comprising genes encoding  
functional isoprenoid enzymatic biosynthetic pathway comprising disrupted  
PT genes such as dead, mreC, and yfH6, useful for producing carotenoid.

XX PS  
Example 3; SEQ ID NO 43; 85pp; English.

XX CC  
The present invention describes a carotenoid overproducing microorganism  
comprising the genes encoding a functional isoprenoid enzymatic  
CC biosynthetic pathway comprising a disrupted gene chosen from dead, mreC,  
and yfH6. Also described is a carotenoid overproducing *Escherichia coli*  
CC microorganism (II) comprising an upper isoprenoid enzymatic biosynthetic  
CC pathway comprising the genes *dxs*, *dxr*, *ygbP*, *ygbB*, *lycB*, *idi*, *ispA*  
CC and *ispB*, a lower isoprenoid enzymatic biosynthetic pathway comprising  
CC the genes *crte*, *crfB*, *crfI*, and *crfJ*, mutations chosen from a mutation in  
CC the *hrs* gene of the 3159 nucleotide sequence of SEQ ID NO:35 (ADQ14658),  
CC mutation in the *rpsA* gene of the 2904 nucleotide sequence of SEQ ID NO:37  
CC (ADQ14660), mutation in the *rpoC* gene of the 5454 nucleotide sequence of  
CC SEQ ID NO:38 (ADQ14661), mutation in the *yfH6* gene of the 1845 nucleotide  
CC sequence of SEQ ID NO:39 (ADQ14662), or mutation in the *rhlA* gene of the  
CC 2676 nucleotide sequence of SEQ ID NO:41 (ADQ14664), where the genes of  
CC the lower isoprenoid enzymatic biosynthetic pathway reside on an  
CC autonomously replicating plasmid comprising a replicon chosen from p15A  
CC and pMB1. (I) chosen from bacteria, yeasts and filamentous fungi such as  
CC *Aspergillus*, *Trichoderma*, *Saccharomyces*, *Pichia*, *Candida*, *Hansenula*,  
CC *Salmonella*, *Bacillus*, *Acinetobacter*, *Zymomonas*, *Agrobacterium*,  
CC *Erythrobacter*, *Chlorobium*, *Chromatium*, *Flavobacterium*, *Cytophaga*,  
CC *Rhodobacter*, *Rhodococcus*, *Streptomyces*, *Brevibacterium*, *Corynebacteria*,  
CC *Mycobacterium*, *Deinococcus*, *Paracoccus*, *Escherichia*, *Erwinia*, *Pantoea*,  
CC *Pseudomonas*, *Sphingomonas*, *Methylobacter*, *Methylobacter*, *Methylococcus*,  
CC *Methylophilus*, *Methylobacterium*, *Methylocystis*, *Alcaligenes*,  
CC *Synechocystis*, *Synechococcus*, *Anabaena*, *Thiobacillus*, *Staphylococcus*,  
CC *Methanobacterium*, *Klebsiella*, and *Myxococcus*, and (II) are useful for  
CC producing carotenoids which involves contacting (I) or (II) with a  
CC fermentable carbon substrate, growing the carotenoid overproducing  
CC microorganism for a sufficient time to produce carotenoid, and optionally  
CC recovering the carotenoid from the carotenoid overproducing  
CC microorganism. The present sequence represents the reporter plasmid  
CC pPCB15 nucleotide sequence, which is used in an example from the present  
XX invention.

SQ Sequence 8609 BP; 1918 A; 2325 C; 2309 G; 2057 T; 0 U; 0 Other;

Query Match 37.4%; Score 2619.6; DB 12; Length 8609;

Best Local Similarity 64.7%; Pred. No. 0;  
Matches 4015; Conservative 0; Mismatches 2169; Indels 24; Gaps 7;  
QY 373 CAGCTGCAGCAAAATTTTACAGCGCATCTTTGAACTTTTACTGCTGCGGACAGCAAAAGC 432  
DB |||||  
DB 1627 CAGTTGCTGGCTGATATCGATAGCCGCTTGTAGTACTGCCGTTCCAGGTGAGCGG 1686  
QY 433 GATCGGTGCGTGGCGGATGCGTGGCGAAGCTGGCGGCGGAGGCAAAAGTATTCCTCT 492  
DB |||||  
DB 1687 GATTGTGGGTGGCGGATGCGTGAAGGACGCTGGCAACCGGCAAAAGTATTCCTCG 1746  
QY 493 TTATTACTGCTGCTGGCAGCGCGATATGGGTGGAGCTGACGCAAAATTCGCTTCTC 552  
DB |||||  
DB 1747 ATGCTGCTGTTATTAACAGCGCGGATCTTGCTGTGCGATCAGTCACGGGGATTTACTG 1806  
QY 553 GATCTGCGCTGTGCAAGTGGAAATGTGCAACGGCGCATCGTATTCGATGATGATTCCTC 612  
DB |||||  
DB 1807 GATTTAGCTGCGGCTTGAATGTGCAATGCTGCTGCTGATTCGTGATGATATGCC 1866  
QY 613 TCGATGGAATACCGCGCATGCGTGGTGGCGGCTTACCGTACCGTACCGAATTTGGTGA 672  
DB |||||  
DB 1867 TGCATGGACGATGCGCAGATGCGTGGGGGCGTCCACCAATTCACACGACGATACGGTGA 1926  
QY 673 AACGTGGCGATTCTCGCGCCATCGCGTGTAGCGCGCATTTGAAAGTGTGATTCGCAAT 732  
DB |||||  
DB 1927 CATGTGGCGATTCTGGCGCGTTCGTTTACTCAGCAAGCGTTTGGGTGATTTGGCGAG 1986  
QY 733 GCACCGCGTTTGGCTGCCATACATAAATCTGAAGCGATTGCTGAATCTTCGCTGCCGCT 792  
DB |||||  
DB 1987 GCTGAAGGTCTGACCGCGATAGCCAAACTCGCGCGTGTGCGAGTGTCCACTGCGATT 2046  
QY 793 GCGCTGAGGGTGTAGTGAAGGCAATTCAGAGTCTGCAGACGCGGACGCGAGACCGCG 852  
DB |||||  
DB 2047 GCGATGCGGGTCTGGTTTCCAGGGCCAGTTTAAAGGACCTCTCGGAAGCGATTAACCCCGC 2106  
QY 853 ACCCGGAGCGATCGCCATGACCAACGAACTGAAACAGCGTGTGTTTTCGCGCACG 912  
DB |||||  
DB 2107 AGCGCGATGCCATCTACTGTAACCAATCAGTTTAAACAGCACGCTGTTTTCGCGTCA 2166  
QY 913 CTGCAAAATGGCGGATTTGCCCTGACCGTTTCAACCGCAGGTGGCGCAAGACCTTAGCTTC 972  
DB |||||  
DB 2167 ACGCAAAATGGCGTCCATTTCGCGCCCAACGCGTCTCGGAAGCGGTGAGAACCTTCGTCATCGT 2226  
QY 973 TTGCGCGAGATTTGGCGCAGCGGTTTAACTGCTGCAGACCTTCGCGCAGGTGCGAAA 1032  
DB |||||  
DB 2227 TTCTCGCTCGATCTGCGCAGCGCTTTTCACTGCTTGCAGATCTTACCGATGCGCAGC 2286  
QY 1033 CACACGGTAAAGATGTGCACAGGATCAGGCAAAATCAGCGTGTGTACAGATGCTCGGT 1092  
DB |||||  
DB 2287 GATACCGGAAGACATCANTCAGGATGAGGTAAATCAACGCTGTCTAATTTATTAGGC 2346  
QY 1093 GCTGACGGCGGAAAGCTGCGCTGCGGATCACCTGCGCAGCGCAGATGCGACCTTGCC 1152  
DB |||||  
DB 2347 TCAGCGCGGTGGAAGACGCTGCGACAGCATTTGCGCTGCGCAGTGAACACCTTTCC 2406  
QY 1153 TCGCGCTGCGCATGCGGCGATCGCGCTGCGCAA---TATATGACCGCGCTGTTTAAATCAA 1209  
DB |||||  
DB 2407 GCGGATGCGCAAAAGCGGCAATTCACACCACTTTTATTTCAGGCTGTTTGGTGA 2466  
QY 1210 CAGCTAGCGATATTCAACTGAGCGGCTCAGCGGTGGCGCACTTTGCGGTGATCGCGC 1269  
DB |||||  
DB 2467 AAATCGCTGCGCTGAGTTAAGATGCTGCA-----TGAGCCATTTTGGGTGATCGCAC 2521  
QY 1270 CGCGCGCTTACAGCACTTTTACCGGTGCGGCGTTAGCAGCAAAACGCTGCTGGCGCGCG 1329  
DB |||||  
DB 2522 CGCGCTTTTTCAGGCAATGTTGCGCTCTGCAAAACCTTTCAGGAATTTAGTGGCGCGCG 2581  
QY 1330 GGCATGCAATCATTATTCAGAGCGCGATGCGCGCATTTGCTTAGGAGCAAGCA 1389  
DB |||||  
DB 2582 GTCATGCTTTTACGTTTTCAGCAACATGACTGCAAAAGCGCTGTGTAACGGGCGAGATA 2641  
QY 1390 TCGATTTTTCGCGTGGCGCAACAGACGATCTCGGTTGCTGGCGCGCGCTGTTGC 1449  
DB |||||

Db 2642 TCGGATTCAGACCGCTCGAGCTGCAAAAGCATCTCCCGTTCTTATCGCACCTGCTGC 2701  
Qy 1450 ATCGGCTGGCTCGCGGCGGCTGTGCTGTTTCGGGTGATCAGCATCTCGGCTCCT 1509  
Db 2702 ACCTGGCGCGCACCATCTCGAGCCCTCGATGTTACGACTGATCAATGAATGGCACGTA 2761  
Qy 1510 GCACCGATATGTTGTCGCGCAATCTGCTCGGTTACTGAAAGCATTTGAACATCGATGGCG 1569  
Db 2762 CCAGCGATATGTTTTCGCGGAATCTGCCCGCTTTTCATGCGTTGAGATAGAGGGG 2821  
Qy 1570 TGATCGCGACGAATGGAAGCGGCGGCGGATGCTGCTGAAGCGGTGCAATCTGCCGT 1629  
Db 2822 TGATCGTTGATCAAAATGGAGCGGCGAGGTGCAAGTAGTCGCAAGAGCGTCAAGTCTGCCGT 2881  
Qy 1630 TTGTTTCGGTGGCTGGCTTGGCGGTCAATCGGTGAAGCGGGATTCGCTTGGCGTGA 1689  
Db 2882 TTGTTTCGGTGGCTGGCGGTGCGGCTCAACCGGAACCGGTTTGCTCTGGCGGTGA 2941  
Qy 1690 TGCCCTTCGTTTTCAGAGGATGACAAAGCGCTGAAACGTTTTCAGGCCAGCAGCGATA 1749  
Db 2942 TGCCCTTCGAGTACGGCACCGGATCGGCTCGGAGCGCTATACCAACGCGAATAA 3001  
Qy 1750 TCTATGATCGCATGCTGTCTACGGCGACGTGATCTCAAAACAGCGGCGGGGTTTA 1809  
Db 3002 TTTATGACTGGCTGATGCGACGTCAAGTCTGATCGCGCATCATGCAATGAGATGG 3061  
Qy 1810 ATTTGACGGAGCGCGGATTTACATCATGCTGCTGCGCGCTGGCACAATAACAGCAGA 1869  
Db 3062 GTTTAGCCCCGGTGAAACCTGCATCATGTTTCTCCACTGGCACAATAACAGCAGT 3121  
Qy 1870 TGTGTCGGCGCTTGTATTTCCAGCTCAGCAACTGCGCCCTGCTATCAAGCGCGTGGGC 1929  
Db 3122 TGATCCCCGAATGGATTTTCCCGCAAGCGCTGCCAGCTGCTTTCATCGGTGGAC 3181  
Qy 1930 CACTCGCG - - - CCGCGTTTCTCTGCGCGCTCAATCGCGCTGGCGGCGCTGCGTGC 1986  
Db 3182 CGTTACGGCAACCCAGGGGAGCGCGGGTCTCAACATCTTATTTTCGCTCCCGGACA 3241  
Qy 1987 AGCGGTGTTTATGCTCGCTGGGTACGCTGCAAGGCCATCGCTTCCGGCTTCTGCG 2046  
Db 3242 AACCCGTAATTTTGCTCGCTGGGACCCCTGGAGGACATGTTATGSCCTGTTTCAGA 3301  
Qy 2047 ATCTGCGCAGGCGTCCGCCAGCTCGGCTATGCTGTGATTCGCCCATGTTGGGGAT 2106  
Db 3302 CCATGCCCAAGCGCTCGAAGAGTGAATGCGAGTTACTGTTGGCACAATGCTGGCGGC 3361  
Qy 2107 TAAACCCGAACAGACGATCAGCTGAGCTCGCTGGCGGCGGTGGGTGACGGAATTCG 2166  
Db 3362 TCTACGACGCGCAGGCGAGTGAATGCGCCGGGCGGGGACATTCAGGTTGTGGATTTG 3421  
Qy 2167 TCGATCAGCGCGCAGCGCTTACAGCAGCGAGCTGTTTATCACTCATGCGGGTTAAACA 2226  
Db 3422 CCGATCAATCCGAGCAGCTTTCAGGCGACAGTTGACATCAACATGTTGGGATGAATA 3481  
Qy 2227 GCGCGCTGAAGCAGCTTGAATCGGTAACCGATGCTGCGCTGCGGATGCTTTTGATC 2286  
Db 3482 CGGTACTGAGCAGTATGCTTCCCGCACACCGCTACTGCGGCTGCGCTGGCATTTGATC 3541  
Qy 2287 AGCCCGGCTGGCGCGCGCATGATGAGCATGAGTGGTGGTGGCGCATCAGCTTTA 2346  
Db 3542 AACCTGGCGTGGCATCACGAATTTGTTATCATGGCATCGGCAAGCGGTGCTTCGGTTTA 3601  
Qy 2347 GCGGTGTTTCACTAACTGGAGCAGCATCTGCAACAGCTGCTGACCGACGATCGTTACGCG 2406  
Db 3602 CTACCAAGCCATGGCTGGCGCGCAGATTCGATCGCTGCTGACTAACCGATTAACCGC 3661  
Qy 2407 TAGCGATGTCAGCGATTCAGGCGCAGCTGACGCGCGAGCGGTTCGCAAGCTGGCGCG 2466  
Db 3662 AGCGTATGACAAAAATTCAGGCGCGATTTGCTTGGCAGCGCGCACACCGCGCGCGC 3721  
Qy 2467 ACATCTGACGAGGCGCTGTGCGCAGCAGCAAGTCTGCTGGCGGAGGCGACCTGATCG 2526  
Db 3722 ATATTGTTGACAGGCGATCGG - GACCTGTGACCCAGTACTCAGTGGCAGGATTTATGCA 3780

Qy 2527 CACCAATACGATGTGATTTTGGTGGTGTGACTGGCGAATGCTGTGATTTGGCTGCG 2586  
Db 3781 ACCGACTATGATCTCATTTCTGGTGGTGGCTGTGGCTAATGGCTTATCGGCTCCG 3840  
Qy 2587 TCTGGTCAATTCAGACCAACAACTGAATGCTGTTGCTGAGAGCGATGCGCATCGGC 2646  
Db 3841 GCTTCAGCAACAGCATCCGATATCGGATCTTCTTATTTAGAGGGGTCTGAGGGGG 3900  
Qy 2647 AGGCAATCATACCTGGTCTGTTTCATCAGCGATCTCAGCGCGCAAACTTCGCTGCT 2706  
Db 3901 AGGGAACCATACCTGGTCTTTCATCAGAGAGATTTACGCTGAATCAGCATCGCTGAT 3960  
Qy 2707 GCAACCGCTGATTTACCGTGGCTTGGTCAAGTTATCAGTGGCTTTTCTGGCTGCGCG 2766  
Db 3961 AGCGCGCTTCTGGTCCATCATCTGCCCGACTACCAAGTTCGTTTCCCGCAACGCGTGC 4020  
Qy 2767 CAATTCGACGGGATTTATGTTTCATTCGATCAGCGATTTTGGCCGCCATCTTACGC 2826  
Db 4021 CCATGTGAACAGTGGCTACTACTGCTGACCTCCCGCATTTTCGCGCGGATCTCCGCA 4080  
Qy 2827 GCGGATGGTCAACGATCTGTGACAAACACAGCGCTACACAGGTAAACCAACAGCT 2886  
Db 4081 ACAGTTTGGACAAATTTATGGCTGCATACCGCGTTTCAGCGTTTCATGCTGAATCGT 4140  
Qy 2887 GACGCTGGCGATGGCGTGAACCTTGTGCGCAAGTGGTGAATGATGTCGCGCTGCA 2946  
Db 4141 CCAGTTACGGATGGCGGATTTATTCATGCCAGTACAGTATCGACGAGCGGGTTACAC 4200  
Qy 2947 GCGGACGACATCTGAGCTGGGTATCAGGTGTTTCTTGGCAAGAGTGGAGCTGCG 3006  
Db 4201 GCCTGATTTCTGCACCTACCGTACGATTCAGGCAATTTATCGTCAAGGATGGCACTGAG 4260  
Qy 3007 GCGCGCGCACGCGCTGACGAGCGGATCTGATGATGCCACCGTGCATCAGCAAGCGG 3066  
Db 4261 GCGCGCGATGGTTTATCGTCAACGATTTATCGATGCGAGCTGATCAGCAAAATGG 4320  
Qy 3067 TTATCGTTTGTCTACGCTGCGCTCAGGCGCGATCGGCTATTTGATTTGAAGATACCA 3126  
Db 4321 CTACCGCTTGTGTTATACCTGCGCTTTCGCAACCGCACTGCTGATCGAAGACACA 4380  
Qy 3127 TTACGTTAACGACCGCGCTGGCGGAGACACCGCTGTCAGCACATCGCGACTATGC 3186  
Db 4381 CTACATTCAGCAAGGCTAATCTTCAGCGCGAACGGCGCTGAGAATCTTCGAGTATGC 4440  
Qy 3187 CAATCAGCAAGGCTGAGCGTACGCTGCTGCTGAAGAGCAGCGATATTAACGAT 3246  
Db 4441 TGCGCGACAGGGTTGCGCTTACAGAGCTTGTGCGGGAAGAACAGGGTGCATTCGCCAT 4500  
Qy 3247 TACCTGAGCGGCAACATCGATCGATTTGCGCAACAGCGCGCGCAAGCTGACGCG 3306  
Db 4501 TACGTTAACGGCGGATATCGTCACTTTTGGCAACAGCAAC - - - CGCAAGCTGTAGCGG 4557  
Qy 3307 CCGCGCGCGGCTGTTTCATGCGCACCGGTTACTCTTGGCTGCGCTGCGCGCTGCGCT 3366  
Db 4558 ATTAAGCGCGCGGCTGTTTTCATCCGCAACCGGCTACTCTTACCGCTGCGCGTGGCT 4617  
Qy 3367 AGCGAGTTGTTAGCAGCGCTGTTGCCCAACCGATCCCTCAGCTCAGCAACATATCGA 3426  
Db 4618 GGCGGATCGTCTCAGCGCGCTGATGTTTACCTCTTCTCTGTTTACAGACGATTCG 4677  
Qy 3427 AGCTTTGCGCGCTGAGTGGCGGACAGCGATTTTTCGCTGCTTAAACCGATGCT 3486  
Db 4678 TCACCTTTGCCAGCAACGTTGGCAACAGGGGTTTTTCCGATGCTGAATCGCATGTT 4737  
Qy 3487 GTTTTGGCGGTAAAGCGCAGCGCTGGCGGTGATGCAACGTTTTTACCGGCTCGA 3546  
Db 4738 GTTTTGGCGGACCGCGCGAGTCACTGCTGGGTGATGAGCAGCGTTCTATGCTTACC 4797  
Qy 3547 TCGCGGTTTAAATGACGCTTTTACGCGGCGCACTGCGCTGCGGATTAACCGGAT 3606  
Db 4798 CGAGGATTTGATGCGCGCTTTTATGCGGGAAAACTCACCGTACCGGATCGCTAGCAT 4857

QY 3607 TCTGTGGCGCAAGCCGCGGTGCCCATCGGTGAAGGGCTGCGGCGCTGTTGAATCTGT 3666  
DB 4858 TCTGAGCGGCAAGCCGCGGTGCCCATCGGTGAAGGGCTGCGGCGCTGTTGAATCTGT 4917  
QY 3667 CGAACCAGGGAAGAAAATGAACCGCACTTATGTGATTGGCGCAGGCTTTGCGGCGCTG 3726  
DB 4918 TCGTTGAAGAGCGACTACATGAACCAACTAGGTAAATTTGGTGGGCTTTGTGGCCCTG 4977  
QY 3727 GCGTGGCGGATTCGCTCGAAGCGGCGGCATACCAACCACTTACTCGAGCAGCGCGAC 3786  
DB 4978 GCACTGGCAATTCGTTTACAGCCGCGAGGTATTCCTGTTTGTCTGCTTGAGCAGCGCGAC 5037  
QY 3787 AAACCGGCGGAGCGGCTATGTGTTGAGGACAGTGGCTTTACCTTCGATCCGAGACCC 3846  
DB 5038 AAGCGGGTGGCGGCGCTTATGTTTATCAGAGCGAGGCTTTACTTTTGTAGCAGGCGCT 5097  
QY 3847 ACGGTGATCACCAGTCCAGCGCCATCGAAGAGTTGTTCACTGCGCAGGAAATCGCTC 3906  
DB 5098 ACGTTATCACCAGTCCAGCGCGATTTGAAGAACTGTTTGCTCTGGCCGCTTAAACAGCTT 5157  
QY 3907 ACGGATTTAGTGGAGCTGATGCGGTAAACCGCTTCTATCGCTGTGCTGGGAAGATGGC 3966  
DB 5158 AAGGATTTAGTGGAGCTGATGCGGTAAACCGCTTCTATCGCTGTGCTGGGAGTCCGGC 5217  
QY 3967 AAACAGCTTGATTAAGCAATTAATCAGCCGCTGCTGGAGCAGAGATCCCGACGTTCAAT 4026  
DB 5218 AAGGCTTTCAATTAAGCAATTAATCAGCCGCTGCTGGAGCAGAGATCCCGACGTTTAA 5277  
QY 4027 CCGCAAGATGAGAGGCTATCGTCAATTTCTTGCTTATTCAGCTGAGTATTTAGAG 4086  
DB 5278 CCGCGGATGTTGCGGCTTATCAGGCTTCTTGACTATTCGCTGCGCTATTCATGAG 5337  
QY 4087 GGTATCTGAACACTCGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4146  
DB 5338 GGTATCTGAACACTCGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5397  
QY 4147 CCGGATTTGGAGGCTGCAAGATGCGGCGAGGCTCTACAGATGTTGGCGAAATTTAT 4206  
DB 5398 CCGGATTTGGAGGCTGCAAGATGCGGCGAGGCTCTACAGATGTTGGCGAAATTTAT 5457  
QY 4207 CAGGAGATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4266  
DB 5458 GAGGATGAGCATCTCTGCGGAGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5517  
QY 4267 TTTGCAACCTCATCTATCTATCTTAAATTCATGCTGCTGAGCGTGAATGGGCGGTGG 4326  
DB 5518 TTTGCAACCTCATCTATCTATCTTAAATTCATGCTGCTGAGCGTGAATGGGCGGTGG 5577  
QY 4327 TTTGCGCGCGGCGCACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4386  
DB 5578 TTTGCGCGCGGCGCACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5637  
QY 4387 GCGGCGAGCTGTTACTGTAATCCGGAAGTGAAGCGCTGCGAAACAGCGCAATCGCAT 4446  
DB 5638 GCGGCGAGCTGTTACTGTAATCCGGAAGTGAAGCGCTGCGAAACAGCGCAATCGCAT 5697  
QY 4447 AGCGGCTTCAAGTTAGAGCGGCGAGCAGCTTCTGATGCGCGCTGCTGCTGCTGCTGCT 4506  
DB 5698 CAGGCGCTGCTGAGTGAAGCGGCGAGCAGCTTCTGATGCGCGCTGCTGCTGCTGCTGCT 5757  
QY 4507 GAGTGGTGCATCTCTACGACAAATCTGCTTGGCGACCATCCGCTGGCAATGAACGTCGG 4566  
DB 5758 GATGTTGTACATACCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5817  
QY 4567 ACATCGCTCAAGGTTAAGCGCATGAGCACTCGCTGTTGTTGTTGTTGTTGTTGTTGTT 4626  
DB 5818 AAAAATCTGAATCCAAAGGTTATGATACTACTGTTGTTGTTGTTGTTGTTGTTGTTGTT 5877  
QY 4627 CAGCGCATGAACAGCTCGCGCACCAACCGCTGCTGTTTGGCGCGGTTATCGTGAAGTTG 4686  
DB 5878 CATCATCAGATCAACTCGCCCATCATACGCTGCTGTTTGGCGCGGCTACCGTGAAGTTG 5937  
QY 4687 ATGATGAGATTTTCAACAGCAGCGCTGGGAGACGATTTTTCATCTTACTGCAACGG 4746

DB 5938 ATTACGAAATTTTAAACCATGATGCTGGCTGAGGATTTTTCGCTTTATTTACACGCA 5997  
QY 4747 CCTCTGAGCAGCGATCTCGCTGCTGCGCCCGCTGCGGCGAGCTTTTATGTTGTTAGCG 4806  
DB 5998 CTTTGTGTACCGATCTCGCTGCTGCGGAAAGGTTGCGGCGAGCTTATTTATGCTGGCG 6057  
QY 4807 CCGGTGCGCATCTCTGCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4866  
DB 6058 CCGTCTTCCACACTTAGGCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6117  
QY 4867 GATCGAAATTTTGTCTTATCTGAGCAGCACTATCATGCTGCTGCTGCTGCTGCTGCTGCT 4926  
DB 6118 GATCGTAATTTTGTCTTATCTGAGCAATTAATCATGCTGCTGCTGCTGCTGCTGCTGCT 6177  
QY 4927 ACACACAGATGTTTACGCGCTTCTGATTTTCCGCAACCATGCTGCTGCTGCTGCTGCTGCT 4986  
DB 6178 ACGCACGTAATGTTTACGCGCTTCTGATTTCCGCGAGCTCAATGCTGCTGCTGCTGCTGCT 6237  
QY 4987 GCGTTTTCGCTGAGCGCATTTTGAACGAAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5046  
DB 6238 GCGTCTCGGTGAACTTATCTGACCCAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6297  
QY 5047 GCGGATATCAGCAATCTCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5106  
DB 6298 AAGCACATGTAATCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6357  
QY 5107 GCGGTGATCGGTTCGCGCAAGCGCACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5166  
DB 6358 GCGGTATGCTGCTGCGGAGCGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6417  
QY 5167 ATCGAAGCTTTTACTTGAAGCAAGTAAACGAAACCATGCTGCTGCTGCTGCTGCTGCTGCT 5226  
DB 6418 AT---ACGTCATTAATCTGATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6474  
QY 5227 CCAACCGCGCGCAAGCTGTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5286  
DB 6475 CCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6534  
QY 5287 GGTGCT 5346  
DB 6535 GGTGCT 6594  
QY 5347 ATGCGCTGCAAGCAGCGGCGACGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5406  
DB 6595 CCGTCTGCGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6654  
QY 5407 ACAGCGGCGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5466  
DB 6655 ACAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6714  
QY 5467 ACCAGCT 5526  
DB 6715 ATGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6774  
QY 5527 ACAGCATTTACGAGCTTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5586  
DB 6775 AAACGCGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6834  
QY 5587 TCGGTTTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5646  
DB 6835 TGGGCT 6894  
QY 5647 GCGATTTAGGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5706  
DB 6895 GCGATCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6954  
QY 5707 AAAATGCT 5766  
DB 6955 AGGTGGCGCT 7014  
QY 5767 CCGTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5826







Db 2107 AGCGCGATGCCATCTAGCTAACCAATCAGTTTAAACACGACCGCTGTTTTGCGCGTCA 2166  
QY 913 CTGCAAAATGGCGCGATTTGCCGCTTGACGCTTACCGCGAGGTGCGGCAAGACTTACGTTTC 972  
Db 2167 ACGCAAAATGGCGCTCCATTGGCGCCAAAGCGCTCTCGAAGCGGTGAGAACCTGCATCGT 2226  
QY 973 TTGCGCCAGGATTTGGCCAGCGGTTTAACTGCTGAGAGCCTCGCCGACGGTTGCMAA 1032  
Db 2227 TTCTCGCTCGATCTCGGCCAGCGCTTTCAGTTGCTTGACGATCTTACCGATGGCATGCC 2286  
QY 1033 CACACGGTAAAGATGTGCACAGGATCAGGCAAAATCCACGCTGTTACAGATGCTCGGT 1092  
Db 2287 GATACCGGCAAGACATCAATCAGATGCAAGTAAATCAACGCTGTCATTTATTAGGC 2346  
QY 1093 GCTGACGGCGGAAAGCTGCGCTTGGCGGATCACTTGCAGCGCGAGATGCACACCTTGGC 1152  
Db 2347 TCAGGCGGCTGGAAGAACGCTGCGACAGCATTTGGCGCTTGGCCAGTGAACACCTTTCC 2406  
QY 1153 TCGGCTTGCATCGGGGATCGCACTCGCAAA-----TATATGACGCGCTGTTTAAATCAA 1209  
Db 2407 GCGGATGCAAAACGCGCATTTCCACACCCAACTTTTATTCAGGCTTGGTTTGACAAA 2466  
QY 1210 CAGCTAGCGATATTCAACTGAGCGGCTCAGCGGTGGCCACTTTGCGGTGATCGCG 1269  
Db 2467 AAATCGCTGCGCTCAGTTAAGATGCTGCA-----TGAGCCATTTTGGCGGTGATCGCAC 2521  
QY 1270 CGCGGCTCTACAGCCACTTTACGCGGTTGACGCGGTTAGCAAAACGCTGCTGGCGCGG 1329  
Db 2522 CGCCCTTTTACGCCATGTTGCGGCTCTGCABAACTTGTCTCAGGAAATAGTGGCCGCG 2581  
QY 1330 GCATCGCATCAATTCATCAGCAAGCGGATGCCGCACTTTGCTTAGCGACGAACGCA 1389  
Db 2582 GTCATCGTTTACGTTTTCAGCAACATGACTGCAAAAGCGCTGTAAACGGCGACGATA 2641  
QY 1390 TCGATTTTGTTCGCTGCGCAACAGACGCTCTGCGGTTGCTGGCGCGCGTTGTC 1449  
Db 2642 TCGGATTCAGACCGTGGACGTGCAAAACGATCTCCCGGTTCTTATCGCACCTTGCTGC 2701  
QY 1450 ATCGGCTGCGCTCGCGCGCGCTGCTGCTGTTGCGGTGATGACAGATCTCGGCTCCT 1509  
Db 2702 ACTTGGCGCGCACCACTCGGACCTCGATGTAGCACTGATCAATGAATGGCACTA 2761  
QY 1510 GACCGATATGCTGTGCCCGAACTGCTGCGGTACTGAAAGCATGGAACATCGATGGCG 1569  
Db 2762 CCAGGATATGCTTTGCCGGAATGCGCGCGCTTTTCATGCTTGAGATAGAGGGCG 2821  
QY 1570 TCATCGCGACGAATGGAGCGCGCGGATGCTCGCTGAGCGCTGCATCTCGCT 1629  
Db 2822 TGATCGTTGATCAAAATGGAGCGCGGAGGTGCGAGTAGTCGCAAGCGTCAAGTCTGCGGT 2881  
QY 1630 TTGTTTGGGTGGCTGCGGCTTTGCGGTTCAATCTGTGAAGCGGGAATTCGCTTGGGTGA 1689  
Db 2882 TTGTTTGGGTGGCTGCGGCTGCGGCTCAACGCGGACCGGTTTTCCTCTGGCGGTGA 2941  
QY 1690 TGCCCTTCCGTTTTCACAGGATGACAAAGCGCTGAAACGTTTTCAGGCGACGACGATA 1749  
Db 2942 TGCCTTTCAGTACGCGCACGCGATGCGGCTCGGGAACGCTATACCAACGCAAAAAA 3001  
QY 1750 TCTATGATCGCATATCGCTGCTCAGCGGACGATGATCTCAACACGCGCGCGGCTTTA 1809  
Db 3002 TTTATGACTGGGTGATGCGACGTCACGATCGTGTGATGCGCATCATGATGCAAGTGG 3061  
QY 1810 ATTTGACGAGCGCGCGGATTTACATCAGTGTGCTGCGGCTGGGCAAAATCAGCCAGA 1869  
Db 3062 GTTTAGCCCGGCTGGAATACTGATCATTTGTTTTCTCCACTGGCAAAATCAGCCAGT 3121  
QY 1870 TGGTCCGCGCTTTGATTTTCCACGTCAGCAACTGCCCGCTGCTATCAACGCGGTGGCG 1929  
Db 3122 TGATCCCGAACTGGATTTTCCCGCAAGCGCTGCCAGACTGCTTTTATGCGGTGGAC 3181  
QY 1930 CACTCGCG-----CCCGGTTTCTCTGCGCGCTCCATGCGCGCTTGGCCAGGCTGCGT 1986

Db 3182 CGTTAGGCAACCCACGAGGACCGCGGGTCATCACTTCTTATTTTCGTCGCCGACA 3241  
QY 1987 AGCGCGTGGTTTANCCCTGCTGGGTACGCTGCAAGGCAATCGCTTCGCGCTGTTCTGC 2046  
Db 3242 AACCCGCTATTTTGGCTCGCTGGGCACTCCCTCAGGACATCGTTATGSCCTGTTCAGGA 3301  
QY 2047 ATCTGGCGAGCGGTGCGCGAGCTGCGGTATCCTGCTGATCGCCCATCTGGGGAT 2106  
Db 3302 CCATCGCCAAAGCCTGCGAAGAGTGGATGCGCAGTTACTGTTGGACACATGCTGGCGGC 3361  
QY 2107 TAAACGCCGAACAGACGATCAGCTGGAGCTCGCTGGCGCGCTGGGTGACGATTTCCG 2166  
Db 3362 TCTCAGCCACGACGAGGTGACTGGCCCGGGCGGGACATTCAGGTTGTGATTTTG 3421  
QY 2167 TCGATCAGCGCGACGCTTACAGACGCGAGCTGTTTATCTACTCATGCGGTTTAAACA 2226  
Db 3422 CGATCAATTCGCGACACTTTTACAGGACAGTTGCAATCAATCACATGCTGGGATGAATA 3481  
QY 2227 GCGCGCTGMAAGCACTGNAATGCGGTACGCGATGCTGGCGCTGCGGATGCTTTGATC 2286  
Db 3482 CGGTACTGACGCTATTGCTTCCGCAACCGCTACTGCGCTGCGCTGGCATTTGATC 3541  
QY 2287 AGCCCGGCTGGCGCGGCAATTGATGGCATGAGTTGGTCCGCGCATCACGCTTTA 2346  
Db 3542 AACCTGGGTGGCATCAAGATTTTATCATGGCATCGGCAAGCGTGGCTCTCGTTTA 3601  
QY 2347 GCGGTGTTTCACTCACTGAGCAGCATCTGCAACAGCTGCTGACCGAAGATCGTTACGCG 2406  
Db 3602 CTACAGCATGCGCTGGCGGCGAGTTGATCGCTGCTGCTGCTAACAACCGATTACCCGC 3661  
QY 2407 TACGATGTACGAGATTACGCGGAGCTGACGCGCGAGCGGTTGCCAGCTGCGCGCG 2466  
Db 3662 AGCGTATGACAAAAATTCAGCGCGCATTTGCGTCTGGCAGCGGCGCACACGAGCGCGCG 3721  
QY 2467 ACATCGTCCAGCAGCGCTGTGCCAGCAGCAAGTCTGCTGGCGAGGCGACTGATGCG 2526  
Db 3722 ATATGTTGMAACAGCGATGCG-GACCTGTACGCCAGTACTCAGTGGGAGGATATGCA 3780  
QY 2527 CACGCAATACGATGTGATTTTGGTGGTGTGAGTGGCGAATGGCTTGAATGCGCTGCG 2586  
Db 3781 ACCGCATATGATCTCATTTGCTGGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3840  
QY 2587 TCTGGTCAATGACGACCACTCAACTGAAATGCTGTTGCTGGAGAGGAGTGCATCCGCG 2646  
Db 3841 GCTTCAGCAACAGCATCCGGATATCGGATCTTGTCTTATGAGCGGGGCTCTGAGCGCG 3900  
QY 2647 AGGCAATCATACTCGTGTGTTTCAACAGCGATCTCAGCGCGCAACAACTTCGCTGGCT 2706  
Db 3901 AGGGAACCATACCTGCTCTTTCAGAGAGAGATTTAAGCTGAATCAGCATCGCTGGAT 3960  
QY 2707 GCAACCGCTGATTAACCGTGGTGGTTCAGGTTTATCAGTGGCTTTTCTGCGCTGCGCG 2766  
Db 3961 AGCGCGCTTGTGGTCCATCACTGCGCGACTACCAAGTTGCTTTCCCGCAACGCGCTCG 4020  
QY 2767 CAATCTGACGGGATTTATGTTCCATGCGCATCAGCGATTTTGGCGGCACTTTTACGC 2826  
Db 4021 CCACTGAACAGTGGCTACTACTGCTGACCTTCCGCGCATTTTCCGCGGATATCTCCGCA 4080  
QY 2827 GSCGATGGGTGACATCTGGCAAAACACAGCGGTACAAAGGTAAACCCACGAGT 2886  
Db 4081 ACAGTTTGAACAACTTTATGCTGATACCGCGTTTTCAGCGTTTCATGCTGAATCGGT 4140  
QY 2887 GACGCTGGCGGATGGCGGTGAATTTGCTGCGCAAGTGGTGAATGATGGTGGCGGCTGCA 2946  
Db 4141 CCAGTTAGCGGATGGCGGATTTATTCATGCGCAGTACAGTGAACGACGGGTTACAC 4200  
QY 2947 GCGGACGCAATCTGACAGCTGGGTTATCAGGTGTTTCTTGGCAAGAGTGGCAGCTGGC 3006  
Db 4201 GCTGATTTGCACTACGCGTAGGATTTCCAGGCAATTTATCGGTGAGGAGTGGCACTGAG 4260  
QY 3007 GCAGCGGACGCGCTGACAGCAGCGATCTGATGATGATGCCACCGCTGATCAGCAGCGG 3066  
Db 4261 CGCGCGCATGTTTATGCTCACCGATATATGATGCGACGCTGATCAGCAAAATGG 4320

QY 3067 TTATCGTTTGTCTACAGCTGCGCTCAGCGCCGATCGGCTATTGATTGAAGTATACCCA 3126  
DB |||||  
QY 4321 CTACCGCTTTGTTTATACCTGCCGCTTTCCGCAACCGCACTGCTGATCGAAGACACACA 4380  
DB |||||  
QY 3127 TTACGTTAAACAGCCGCGCTGGCGAGAACACCGCTCGTCAGCACATCGCCGACTATGC 3186  
DB |||||  
QY 4381 CTACATTGACAGGCTAATCTTCAGGCCGAACGGGGCGTCAGAACATTCGGATTATGC 4440  
DB |||||  
QY 3187 CAATCAGCAAGGCTGGACCTGAGTACGCTGCTGCTGAGAGAGCAGGCAATATTACCGAT 3246  
DB |||||  
QY 4441 TCGCGCACAGGGTTGGCCGTTACAGACGTTGCTCGGGAGAAACAGGGTGCAATTCGCCAT 4500  
DB |||||  
QY 3247 TACCTTGAGCGCAACATCGATCTTGGAACAGCAGCGCGCCGAAGCGTGACAGCG 3306  
DB |||||  
QY 4501 TACGTTAACGGCGGATAATCGTCAGTTTGGCAACAGCAAC---CGCAAGCCTGTAGCGG 4557  
DB |||||  
QY 3307 CCTGGCGCGCGGCTGTTTCATGCCCACACCGGTTACTCTTGGCGCTCGCGCTGGCGCT 3366  
DB |||||  
QY 4558 ATTACGGCGCGGCTGTTTCATCCGACAAACCGCTACTCCCTACCGCTCGCGGTGGCGT 4617  
DB |||||  
QY 3367 AGCGGATGTTGAGCAGCGCTGTGCCCACCGATGCCCTCAGCTCAGCCAAACATATGA 3426  
DB |||||  
QY 4618 GGCCGATCTCTCAGCGCGCTGATGTGTTTACCTCTTCTCTGTTCACAGACGATTGC 4677  
DB |||||  
QY 3427 ACGCTTGGCCGCTCAGCAGTGCGGCAACAGCGATTTTTCCGCTGCTGCTAAACCGCATGCT 3486  
DB |||||  
QY 4678 TCACITTTGCCAGCAACGTTGGCAGCAACAGGGTTTTTCCGCATGCTGATCGCATGTT 4737  
DB |||||  
QY 3487 GTTTTTGGCGGTAAAGCCGACAGCGCTGGCGGTGATGCAACGTTTTTACCGCTCGA 3546  
DB |||||  
QY 4738 GTTTTTAGCGGACCGCGCGAGTCACGCTGGCGTGTGATGACAGCGTTTCTATGGCTTACC 4797  
DB |||||  
QY 3547 TGC CGGTTAAATTAGCGCTTTTACCGGGCAACTGCGCCCTGCGGGATAAACCGGAT 3606  
DB |||||  
QY 4798 CGAGGATTTGATTGCGCGCTTTTATGCGGGAAACCTCACCGTGACCGATCGGCTACGCAT 4857  
DB |||||  
QY 3607 TCTGTGCGCAAGCGCGCTGCCATCGTGAAGCGCTGCGCGCTGTGCAATCTCTGT 3666  
DB |||||  
QY 4858 TCTGAGCGCAAGCGCGCTTCCGTTCCGCTTTTCCGGGATTTGAGGCAATTAAGCACTCA 4917  
DB |||||  
QY 3667 CGAACAGGAGAAAGAAATGAACCACTTATGTGATTGGCGCAGGCTTTGGCGGCTG 3726  
DB |||||  
QY 4918 TCGTTGAAGAGCGACTACATGAAACCACTACGTTAAATTGCTGCGGCTTTGCGGCTG 4977  
DB |||||  
QY 3727 CGCGTGGCATTCGCTGCAAGCGCGGCGATACCAACCTTACTCGAGCAGCGCGAC 3786  
DB |||||  
QY 4978 GCACTGGCAATTGGTTTACAGGCGCGCAGGTAATCTCTGTTTGTCTTGAAGCAGCGCGAC 5037  
DB |||||  
QY 3787 AAACCGGCGGACCGCGCTATGTTTGAAGCAGTGGCTTTTACCTTCGATGCGCGACCC 3846  
DB |||||  
QY 5038 AAGCGGTTGGCGGCTTATGTTTATCAGGACGGGCTTTACTTTGATGACGGCCT 5097  
DB |||||  
QY 3847 ACGGTGATCAACGATCCCAGCGCCATCGAAGAGTGTTCACGCTGGCAGAAATCGCTC 3906  
DB |||||  
QY 5098 ACCGTTATCACCGATCCCAGCGGATTGAAGAACTGTTGCTCTGGCGGTAACAGCTT 5157  
DB |||||  
QY 3907 AGCGATTACGTCAGCTGATCGGTAACCGCCTTCTATCGCTGTGCTGGGAGATGCG 3966  
DB |||||  
QY 5158 AAGGATTACGTCAGCTGTGGCGGTCAGCGCTTTTATCGCTGTGCTGGGAGTCCGGC 5217  
DB |||||  
QY 3967 AAACAGCTTGATTACGCAATATACGCGCTCTCGAGCAGCAGATCGCCAGTTTCAAT 4026  
DB |||||  
QY 5218 AAGGCTTCAATTACGATAACGACACGAGCCAGTTAGAGCGCAGATACAGCAGTTTAA 5277  
DB |||||  
QY 4027 CCGCAAGATGAGAGGCTATCGTCAATTTCTTTGGCTATTCAGCTGAAGTATTTAGAGAG 4086  
DB |||||  
QY 5278 CCGCGGATTTGCGGGTTATCGAGGCTTCTTTGACTTATTCGCTGCGGCTTCAATGAG 5337  
DB |||||  
QY 4087 GGTATCTGAAACTCGSCAGGTCGCTTCTGACGTGCTGACATGCTGCGCGTCGCG 4146  
DB |||||  
QY 5338 GGCTATCTGAAGCTCGCACTGTCCTTTTATCGTTCAAGACATGCTTCGGGCGCGG 5397  
DB |||||

QY 4147 CCGCAGTTGGACGCTCTCGAAGCATGGCGCAGCGTCTACAGCATGGTGGCGAAATTTATT 4206  
DB |||||  
QY 5398 CCCAGTTGGCAAGCTGCGAGCATGGCGAGCGTTTACAGTAAAGTTGCGCGCTACATT 5457  
DB |||||  
QY 4207 CAGGAGATCATCTGCGTCAAGCGGTTTTCTTCCATCATTTGCTGGTGGGCGGTAATCCT 4266  
DB |||||  
QY 5458 GAGGATGAGCATCTTCGCGAGCGGTTTTCTTTCTACTCGCTCTTAGTGGGGGGAATCCG 5517  
DB |||||  
QY 4267 TTTCCAAAGCTCATCGATCTATACCTTAATTCATCGCTGGAGCGTGAATGGGCGTGTGG 4326  
DB |||||  
QY 5518 TTTGCAACCTCGTCAATTTATACGCTGATTCACGCTTAGAACGGGAATGGGCGTCTGG 5577  
DB |||||  
QY 4327 TTTCCGCGCGCGGACCGCGCGCTGCTGTCAGGGCATGGCGCACTGTTTCGAGCACTTG 4386  
DB |||||  
QY 5578 TTTCCACGCGGTGGAAACCGGTGCGCTGCTCAATGGCATGATCAAGCTGTTTCAAGATCTG 5637  
DB |||||  
QY 4387 GCGCGCAGCTGTTACTGTAATGCCGAAGTGAAGCAGCTGGAACCCAGCGGCAATCGCAT 4446  
DB |||||  
QY 5638 GCGCGCAAGTCTGCTTTAACCGCGGCTCAGTCATATGGAACCCGTTGGGGCAAGAT 5697  
DB |||||  
QY 4447 AGCGCGTTCAGTTAGAGGGGCGAGCGCTTCGATGCCGCGCTGTGGCCTCCAATGCC 4506  
DB |||||  
QY 5698 CAGCGCTGTCAGTTGGAAGACGCGCAGACGCTTGAACCTCGCGGTGGCGTCGACGCT 5757  
DB |||||  
QY 4507 GACGTGTGCATACCTACACAAACTGCTTCGCCCAACCATCCGCTGGCAATGAACGCTGCG 4566  
DB |||||  
QY 5758 GATGTTGTACATACCTATCGCGATCTGCTGTCTCAGCATCCCGAGCGCTAAAGCAGCG 5817  
DB |||||  
QY 4567 ACATCGCTGAAGCGTAAGCGCATGAGCAACTCGCTGTTTGTACTATTTTGGGCTGAAT 4626  
DB |||||  
QY 5818 AAAAACTGCAATCCAGCGTATGAGTAACCTCACTGTTGTACTATTTTGGTCTCAAC 5877  
DB |||||  
QY 4627 CAGCGCATGAACAGCTCGCGCACACACCTGCTGTTTGGCGCGCTTATCGTGAGTTG 4686  
DB |||||  
QY 5878 CATCATCAGATCACTCGCCCATCATACCTGCTGTTTGGGCCACGCTACCGTGAACG 5937  
DB |||||  
QY 4687 ATCGATGAGATTTTCAACAGCAGCAGCTGGGAGACGATTTTTCATCTTACCTGACGCG 4746  
DB |||||  
QY 5938 ATTACAGAAAATTTTAAACCATGATGCTGCTGAGGATTTTTCGCTTATTTTACACGA 5997  
DB |||||  
QY 4747 CCCTGACAGCGGATCCGCTGCTGCGACCGCGCTGCGCAGCTTTTATGTGTTAGCG 4806  
DB |||||  
QY 5998 CTTGTGTACGAGATCCGCTCACTGGCACCGGAAGGTCGGCAGCTATTTATGTCTGGCG 6057  
DB |||||  
QY 4807 CCGGTGCGGCATCTCGGCACCGCTGACATCGATGGGCAACAGGAAGACCOCGCTTGGCG 4866  
DB |||||  
QY 6058 CTTGTTCCACTTTAGSCACGCGCAACTCTGACTGGGCGGTAGAAGGACCCGACTGCGC 6117  
DB |||||  
QY 4867 GATCGAATTTTGTCTTATCTGGAGCAGCACTACATGCGGGATTAAGTCTGAGCAATAGTG 4926  
DB |||||  
QY 6118 GATCGTATTTTGTACTACTCTTGAACAATTAACATGCTGCTGCGTTCGGAAGCCAGTTGGTG 6177  
DB |||||  
QY 4927 ACACACAGAAATGTTTACCGCTTTGATTTTCGCGACACGCTGATGCCCATCACGGCTCG 4986  
DB |||||  
QY 6178 AGCAGCGTATGTTTACCGCTTTCGATTTTCGCGACGAGCTCAATGCTTGGCAAGGTTGCG 6237  
DB |||||  
QY 4987 GCGTTTTTGTGAGCGGATTTTGAAGCGCTGCTGTTCCGCCCGCATAACCGCGAT 5046  
DB |||||  
QY 6238 GCTTCTCGTTGAACCTTATCTGACCCAGAGCGCTTGGTTCGACACCAATAACCGCAT 6297  
DB |||||  
QY 5047 GCCGATATCAGCAATCTCTATCTGTGGTGGCGGTACGCAATCCAGCGCGCGGCTGCC 5106  
DB |||||  
QY 6298 AAGCACTTGAATAATCTTTATCTGTGGCGCAGCACCCATCTCTGGCGGGGCAATCCC 6357  
DB |||||  
QY 5107 GCGGTGATCGGTTCCGCCAAGGCCACCGCAGCTGATGCTGAGGATCGCGCCGAATGA 5166  
DB |||||  
QY 6358 GCGGTAATCGGCTCGCGAGGCGCAGCGCTTAATGCTGGAGGACCTGATTTGACGA 6417  
DB |||||  
QY 5167 ATCGACAGCTTTTACTTGAAGTAAGTAAACGCAACCATGGCGGTGGCTCGAAGAGTTTCG 5226  
DB |||||  
QY 6418 AT---ACGTCATTACTGAATCATCGCTGCAAAACCATGGCGGTTGGCTCGAAAGGTTTG 6474  
DB |||||  
QY 5227 CCACCGCGCCCAAGCTGTTTGTATGACCCGACCGCGCGCAGCAGCTGATGCTGTATGCGT 5286  
DB |||||

Db 6475 CGACTGCATCGACGCTTTTCGACGCCAAACCCGTCGACGCTGCTGATGCTTTACGCAAT 6534  
QY 5287 GGTGTGCTCACTGCGATGATGATTTGATGGCAAAACGCTGGCGGAAGCGGCACGACG 5346  
Db 6535 GGTGCGCGCACTCGGACGACGCTCATTTGACGATCAAAACACTGGGCTTTTCATGCGCACG 6594  
QY 5347 ATGCGGTGCAAGACGCGGACGACGATATGACGATCTGCAAAATGAAACCGCGCGGCT 5406  
Db 6595 CCTTTCGAGATGCTGAGCAGCGCTTCGACGATCTGGAAGCTTTGCAATGAAACCGCTCAGGCT 6654  
QY 5407 ACAGCGCGCGCACATGATGAACCGCGCTTTAGGGCGCTTCAGGAAGTGGCGATCATTC 5466  
Db 6655 AGCGGTTGCAATGACGACGCGCTTTTGGCGGCTTTTGGAGGTCGCGATGGCGC 6714  
QY 5467 ACCAGCTGCGCAACAACTGGGCTTTGATCATCTGGAAGCTTCGCTATGATGCAAGCA 5526  
Db 6715 ATGATATCGCTCCGCGCTACGGGTTTCGACCATCTGGAAGCTTTTGCCATGGATGTGGCG 6774  
QY 5527 ACGNACATTCACGGAGCTTCGATGACACGCTGCTGCTACTGCTATCAGCTCGCGGGGCTGG 5586  
Db 6775 AAACGCGCTACCTGACACATGGAAGCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6834  
QY 5587 TCGGCTTGTATGATGGCGCGCTAAATGGCGGTGCGCGACGAAGCGGTGCTGATCAGGCT 5646  
Db 6835 TGGGCTGATGATGGCGCAATATATGGGCTTCGCGATTAACGCGCTGCTGATCGGCT 6894  
QY 5647 GCGATTTAGGACTGGCGTTTCCAGCTCATTAACATTTGCGCGCACATTTGTAAGATGCGG 5706  
Db 6895 GCGATCTCGGCTGGCTTTCCAGTTGACCAACATTTGCGCGTGATATTTGCGAGATGCTC 6954  
QY 5707 AAATGGTCTGCTATCTGCGCAATCTGCGTATGATGAGGGGATTAACGGCGGATA 5766  
Db 6955 AGTGGGCGCGCTGTTATCTGCTGCTGAAAGCTGCTGGAAGAGAAAGACTGACGAAAGCGA 7014  
QY 5767 CGCTGACTCACCGCAACATCTGACGCTGCTGCGCTCACTGCGAGCGGCTTTAGTGGCG 5826  
Db 7015 ATATGCTGCGCAGAAACCGGAGCGCTTAAGCGGTATCGCGGCGACTGTTACGGG 7074  
QY 5827 AGGCGGAACCTATTATCACTCGGCGGATTCGCGTTTACCGGTTTACCGCTCGGCTCGG 5886  
Db 7075 AAGCGGAACCTATTACGTATCATCAATGGCGGCTGCGCACAAATTTACCTTTACGCTCGG 7134  
QY 5887 CGTGGGCAATCGTACGCTCGGGGCTTTATCGGAAATTTGGCGTCAAAAGTTTACGACG 5946  
Db 7135 CCTGGGCAATCGCAGACGGAAGCAGGTTGACCGTAAATTTGCGGTGAAAGTTGAACAGG 7194  
QY 5947 CCGGTGTGACGCTCGGATTCAGGCGAGCGACAGTAAAGTGAAGTGAAGTGAAGTGAAG 6006  
Db 7195 CCGGTAAGCAGGCTGGGATCATGCGCAGTCCAGCTCCACGCGGAAATTTAAGCTTT 7254  
QY 6007 TGGTGAAGGGGACGTTTGGGATCATCTTCGCGTGTGTCTGCTCTGAAACCGCGTCCGG 6066  
Db 7255 TGCTGACGCGCATCGCTCAGGACGTTTACTTCCGGATGAAGACGATATCCACCCGCTCGT 7314  
QY 6067 CTGCTGTGGAGCGCTCTGTTGATTTTACGTCCGTGACGCTGGCGAGCGTGGCTTG 6126  
Db 7315 CTCTCTCTGGCAGCGCCC-----GATCTAGCGCATCGCTTTCTCTCAGCGTGGCTCG 7368  
QY 6127 CAGCTTATTTCAGCGGTGGCGGTAGAGGAAACCAACAGCACGACGCTTTACGCGCGG 6186  
Db 7369 AAGTTTAGATACGTTGGCGGCTACAGAAACCAAGAGCACGACGCGCTCTTTTCCCT 7428  
QY 6187 CACCGCATGATGCTGCGTGGCCATGTATAGCGCTTTAAGATAGCTTTTGGCGGGAT 6246  
Db 7429 TACAGCATGATCATACGTTGGCCATGTATAACCGTTTCAGTAGCTTTTGGCGGAT 7488  
QY 6247 ATAGCGGAACGCGACGCTTGTATGACACGAGCCATCGTGCAACATGAAGTAGAGCGCGC 6306  
Db 7489 GTAGCGGAACGCGACGCTGTTGTACCACTCGTGGACCAATATAATACAGTAAAC 7548  
QY 6307 GTACGCTGCTCATTCGCGACCACTCCACTGCGAGCGGCGCATGCTTGACACCGACATA 6366

Db 7549 ATAAGCGGTCTGCTGCACCAATCCACTGGAGCGCCAGATTCCTGTACTGCCGAGTA 7608  
QY 6367 AATAGCACAATCGCAGGTACCGCAAAACACCAACCATAAAGATCGTTGAGTCAAACTT 6426  
Db 7609 AATCAGGGCAATCGACACAATGGCGAATACCAACGCGCATAGAGATCGTTAACTTCAAATGC 7668  
QY 6427 ACCGCTGTGCGGTTTCATGTCGACAGATGCCAGCCCATCCCAACCGTGCATGATGTA 6486  
Db 7669 GCGTTTACGCGGTTTCATGATGTGAAGATGCCAGCCCCAACCCAGCGCTGATGATGTA 7728  
QY 6487 TTTATGCGACAGCGCGCTTACGATTTCCATCAACCAACCGTTGCCAACAGATAGCAC 6546  
Db 7729 TTTATGTGCCAGTGCAGCAACCACTTCCATGCGGACCGTGCACAAACACGATCAGGCG 7788  
QY 6547 GTTCCATTAACAGAGCATTTGTCGTCCA 6574  
Db 7789 ATTCCAAATCCACAACATAAATTTCTCAA 7816

RESULT 4  
ADQ77262

ID ADQ77262 standard; DNA; 8609 BP.

AC ADQ77262;

DT 23-SEP-2004 (first entry)

XX Nucleotide sequence of plasmid pPCB15.

KW carotenoid; carotenoid enzymatic biosynthetic pathway; dxs gene;

KW D-1-deoxyxylulose 5-phosphate gene; idi gene; ygbBP gene; ispB gene;

KW yjer gene; oligo-ribonuclease gene; dxs gene; dxr gene;

KW DOXP reductoisomerase; ygbP gene; ispB gene; ychB gene; ispE gene;

KW ygbB gene; ispF gene; lycB gene; ispH gene; idi gene;

KW isopentenyl diphosphate isomerase; ispA gene;

KW farnesyl pyrophosphate synthase; ispB gene;

KW octaprenyl diphosphate synthase; crtB gene;

KW geranylgeranyl pyrophosphate synthase; crtB gene; phytoene synthase;

KW crtI gene; phytoene dehydrogenase; crtX gene;

KW zeaxanthin glucosyl transferase; crtZ gene; beta-carotene hydroxylase;

KW crtY gene; lycopene cyclase; enzyme; ss.

XX Synthetic.

XX WO2004056975-A2.

FN 08-JUL-2004.

XX 19-DEC-2003; 2003WO-US041812.

XX 19-DEC-2002; 2002US-0434618P.

PR (DUPO ) DU PONT DE NEMOURS &amp; CO E I.

XX Cheng Q, Rouviere PE, Suh W;

XX WPI; 2004-525437/50.

XX Novel carotenoid overproducing bacteria such as Escherichia coli or

PT Flavobacterium, comprising genes encoding functional carotenoid enzymatic

PT biosynthetic pathway, useful for producing carotenoids such as

PT astaxanthin and beta-carotene.

XX Disclosure; SEQ ID NO 64; 137pp; English.

XX The specification describes a carotenoid overproducing bacteria, which

CC comprise genes encoding a functional carotenoid enzymatic biosynthetic

CC pathway, where the dxs (D-1-deoxyxylulose 5-phosphate gene), idi, ygbBP

CC and ispB genes are overexpressed, and the yjer gene (oligo-ribonuclease

CC gene) is down regulated. In bacteria of the invention, the carotenoid

CC enzymatic biosynthetic pathway consists of the genes dxs, dxr (DOXP

CC reductoisomerase), ygbP (also known as ispD), ychB (also known as ispE),

CC ygbB (also known as ispF), lycB (also known as ispH), idi (isopentenyl

CC diphosphate isomerase), ispA (farnesyl pyrophosphate synthase), ispB  
CC (octaprenyl diphosphate synthase), crtB (geranylgeranyl pyrophosphate  
CC synthase), crtB (phytoene synthase), crtI (phytoene dehydrogenase), crtX  
CC (zeaxanthin glucosyl transferase), crtZ (beta-carotene hydroxylase) and  
CC crtY (lycopene cyclase). The lvtB and dnr gene are optionally  
CC overexpressed. Carotenoid overproducing bacteria of the invention are  
CC useful for producing carotenoids. The present sequence represents a  
CC plasmid pPCB15, which is used during construction of bacteria of the  
XX invention.

SQ Sequence 8609 BP; 1918 A; 2325 C; 2309 G; 2057 T; 0 U; 0 Other;

Query Match 37.4%; Score 2619.6; DB 12; Length 8609;  
Best Local Similarity 64.7%; Pred. No. 0;  
Matches 4015; Conservative 0; Mismatches 2169; Indels 24; Gaps 7;

QY 373 CAGCTGACGAAATTTCAGCGGATCTTTGAACATTTTACTGCTGCGGACAGCAAGC 432  
DB 1627 CAGTGTCTGGTGATATCATAGACGCTTGTATCAGTTACTGCGGTTACGGGTGAGCGG 1686

QY 433 GATCGCGTGGTCCGCGATCGGTGCGGAACGCTGGGCGAGGCAAGATTCGTCT 492  
DB 1687 GATTGTGTGGTCCGCGATGCTGAAGGACGCTGGCACCGGCAACGATTTCGTCCG 1746

QY 493 TTATTACTGCTGTGCGACGCGCATATGGTGTGAGCTGACGCAAAATGGCGTTCTC 552  
DB 1747 ATGCTGCTGTTATTAAACGCGCGCATCTTGCTGTGCGATCAGTCACGGGGATTA 1806

QY 553 GATCTCGCTGTGCAATGGAATGTTGCAACGCGCATCGCTGATTCGTGATGACATGCC 612  
DB 1807 GATTAGCTGCGCGTGTGAATGTTGATGCTGCTGCTGCTGATTCGTGATATGCC 1866

QY 613 TCGATGATTAACGCGCAGATCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 672  
DB 1867 TGCATGAGCAGTGCAGATGCGTGGGGCGTCCACCATTCACACGAGTACGCGTAA 1926

QY 673 AAGCTGCGGATCTCGCGGCAATCGCGCTGCTTAGCGCGCATTTGAAGTGAATGCCATT 732  
DB 1927 CATGTGCGGATCTGCGGCGGTGCTTTACTCAGCAAGCGTTTGGGTGATTCGCGAG 1986

QY 733 GCACCGGTTGCTGCTGCTATCAATAATCTGAAGCATGCTGTAATCTCCGCTGCGCTC 792  
DB 1987 GCTGAAGGTCTGACGCGCATAGCAAAACTCGCGCGGTGTGCGAGCTGTCCAATCGATT 2046

QY 793 GGCCTGCGGCTTAGTGAAGGCAATTCAGATCTGCAACGCGCAACGAGCGCCG 852  
DB 2047 GGCATGAGGCTGTGTTAGGCGCATTTAAGACCTCTCGAAGCGCATAAACCCG 2106

QY 853 AGCCCGAAGCGATCGCATGACCAACGAACTGAAACACGAGTGTGTTTGGCGCACG 912  
DB 2107 AGCGCGGATGCCATCTGCTTAAACATCAGTTTAAACACGACGCTGTTTGGCGTCA 2166

QY 913 CTGCAATGCGGGATTCGCTGAGCTTACCGCAGTGGCGGCAAGACATTAGCTTC 972  
DB 2167 ACGCAATGCGCTCATTTGCGGCAACGCTCTGCGAAGCGGTGAGAACCTTGCATCGT 2226

QY 973 TTGCGCAGGATTTGGCGCAGGCTTTCAACTGCTGCAACGACCTCGCGCAGGTTGCAAA 1032  
DB 2227 TTCTCGTCTGATCTCGCGCAGGCTTTTCACTGCTTGGACGATCTTACCGATGCGATGCC 2286

QY 1033 CACACCGGTAAAGTGTGCAACGAGATCAGGGCAATTCACGCTGGTACAGATGCTCGGT 1092  
DB 2287 GATACCGGCAAGACATCAATCAGGATGCAAGTAAATCAACGCTGGTCAATTTATTAGGC 2346

QY 1093 GCTGACCGCGGAACTGCGCTGCGGATCACTCGCGAGCGCAGATGACACCTTGCC 1152  
DB 2347 TCAGCGCGGTGCAAGAACGCTTCGACAGATTTGGCGCTGGCGCAGTGAACCTTTC 2406

QY 1153 TGCCTCTGCCATCGCGCATGCGCACTCGCCAA---TATTGACGCGCTGTTTAAATCAA 1209  
DB 2407 GCGGCATGCCAAACCGGCATTCCACCACCACTTTTATTTCAGGCTGTTTGAACA 2466

QY 1210 CAGCTAGGATATTCAACTGAGCGGCTCAGCGGTGGGCCACTTTTGGCTGATCGCG 1269

DB 2467 AACTCGCTCGCTCAGTTAAGGATGCTGCA-----TGAGCCATTTTGGGTGATCGCAC 2521

QY 1270 CGCGCTCTACAGCACTTTTACCGGTTGACAGCGTTAGCACAACGCTGCTCGCGCGCG 1329

DB 2522 CGCCCTTTTTCAGCCATGTTTCGGGCTCTGCAAAACCTTGTCTAGGAATTTAGTGGCGCG 2581

QY 1330 GCATCGCATCACATTCATCCACAGCCGATCCCGCACTTTTGTCTTAGCGAGCAACGCA 1389

DB 2582 GTATCGTGTGTACGTTTTTTCAGCAACATGACTGCAAGCGCTGGTAAACGGGACGATA 2641

QY 1390 TCGATTTTGTGCGCTCGGCCAAACAGACGATCTCTGCGGTTTCGCTGGCGCGCTGTGTC 1449

DB 2642 TCGATTTCCAGACGCTCGGACTGCAACGCACTCTCCCGGTTCTTATTCGACATGCTGTC 2701

QY 1450 ATCGGCTGGCTCGCGCGGCGCTGCTGCTGTTTTCGCTGATCGAGTCAATGAATGACGTA 2761

DB 2702 ACCTGGCGCGCACTCGGACCTCGATGTTTACGACTGATCAATGAATGACGTA 2761

QY 1510 GCACGATATGCTGTGCGCGCAACTGCTCGGTTACTGAAAGCATTTGAACATCGATGGCG 1569

DB 2762 CCAGGATATGCTTTGCGGGAATGCTCCCGCGCTTTTCATGCTTGCAGATAGAGGCG 2821

QY 1570 TGATGCGCGCAAAATGGAAGCGCGGCGGATGCTGCTGAGAGCGCTGCATGTCGCT 1629

DB 2822 TGATGCTTGTCAAAATGGAGCGCGAGTGCAGTAGTTCGAGAGCGCTCAGGTTCTGCCGT 2881

QY 1630 TTGTTTTCGCTGCGCTGCGCTTTCGCTGCAATCTGAAAGCGGATTTCCGTTTTCGCTGTA 1689

DB 2882 TTGTTTTCGCTGCGCTGCGCTGCGCTCAACCGCAACCGGTTTTCGCTTTCGCTGTA 2941

QY 1690 TGCCCTTCGCTTTTTCGCAAGGATGCAAAAGCGCTGAAAGCTTTTTCAGGCGCAGAGCGATA 1749

DB 2942 TGCTTTTCGAGTACGCGCAACAGCGATGCGCTCGGAAACGCTATACCAACGAGCAAAAA 3001

QY 1750 TCTATGATCGCATCATGCTGCTGCAACGCGAGCGGATCTTCAACACGCGCGGCTTTA 1809

DB 3002 TTTATGATGCTGATGCGAGCTCAGGATCGTGTGATGCGGCATCATGCTGAGATGG 3061

QY 1810 ATTTGACGCGCGCGCGGATTATCATGTCCTGTGCGCGCTGGCAAAATACGCGACA 1869

DB 3062 GTTTAGCCCCGGTGAATAAATCTGCATCATTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3121

QY 1870 TGGTGGCGGCTTTGATTTTTCAGTCAAGCAATGCGCGCTGCTATCACCGCTGGCGC 1929

DB 3122 TGATCCCGAACTGGAATTTTCCCGCAAGCGCTGCCAGACTGCTTTCTATCGGTTGAC 3181

QY 1930 CACTCCGCG---CCCGGTTTCTCTGCGCGCTCCATGCGCCCTTGGCCAGCGCTGCTC 1986

DB 3182 CGTTAAGCAACCCAGGGAGCGCGGCTCATCAACTTCTTATTTTCTTCTTCTTCTTCTTCT 3241

QY 1987 AGCCCGTGTGTTATGCTCGCTGGGTACGCTGCAAGGCAATCGCTTCCGCTGCTTCTG 2046

DB 3242 AACCCGCTATTTTTCCTCGCTGGGCACTTGCAGGCACTCGTTATGCGCTGTTTTCAGGA 3301

QY 2047 ATCTGCGCAGCGCTGCGCGCAGCTGCGCTGCTGCTGATCGCCCATTTGTGGGGAT 2106

DB 3302 CCATCGCAAGCTTTCGCAAGGTTGATGCGAGTTACTGTTGCACTCATGTCGCGCC 3361

QY 2107 TAAACGCGCAACAGACGATCAGCTGGAGCTGCTGGCGCGCGCTGGGTGAGCGATTTCG 2166

DB 3362 TCTCAGCACCGCAGGAGTGAATGCGCGCGGGGAGCAATTGAGTTTGGATTTTG 3421

QY 2167 TCGATCAGCGCGCGCTTACAGCAGCGGCTGTTTATCATCTCATCGCGGTTTAAACA 2226

DB 3422 CCGATCAATTCGCGACACTTTCACAGGCACTTGCAGATCACATGTTGGGATGAATA 3481

QY 2227 GCGCTGTGAAGCACTGGAATGCGGTAGCCGATGCTGGCGCTGCCGATTTGCTTTGATC 2286

DB 3482 CGGTACTGAGCGCTATTTCTTCCCGCACACCGCTACTGCGCTGCGCTGGCATTTGATC 3541

QY 2287 AGCCCGGCTGCGCGCGCATTTAGTGGCATGAGTTGGTTCGCGCGCATCACGCTTTA 2346

Db 3542 AACCTGGCGTCATCACGAATTGTTTATCATGGCATCGCAAGCGTGGCTCTCGGTTTA 3601  
Qy 2347 GCGTGTTTATCAATCGAGACAGCATCTGCAACAGCTGTGACGACGATCGTTACGCGC 2406  
Db 3602 CTACACGCAATGCGCTGGCGCGCAGATTGATCGCTGCTGACTAACAACGATTAACCGC 3661  
Qy 2407 TACGGATGTCAGGATTCAGGCGCAGCTCAGCGCGCAGCGGTTGCCAGCGTGCAGCG 2466  
Db 3662 AGCGTATGACAAAATTCAGGCGCGCAATGCGTCTGGCAGCGCGCACACAGCGCGCGCG 3721  
Qy 2467 ACATCGTCGAGCAGGCGCTGTGACAGCAGCAAGTCTGTGGCGGAGCGACCTGATGCG 2526  
Db 3722 ATATTGTTGACAGGCGCATGCG-GACTGTGACGAGTACTCAGTGGCGCAGATTATGCA 3780  
Qy 2527 CACGCAATACGATGTTGTTGCTGGTCTGAGCTGGCGAATGCTTGAATGCGCTGCG 2586  
Db 3781 ACCGCACTATGATCTCTATCTGCTGGTGGCGGCTTATGCGCTTATGCGGCTCG 3840  
Qy 2587 TCTGGCTCAATTGCAACCACTGAATGAATGCCCTGTTGCTGGAGAGCGATCGCGATCGGC 2646  
Db 3841 GCTTCAGCAACAGCATCCGGATATGCGGATCTTGGCTTATGAGGCGGGTCTTGAGCGGG 3900  
Qy 2647 AGCAATCATACCTGCTGTTTATCACAGCGATCTCAGCGCGAACAACCTTCGCTGGCT 2706  
Db 3901 AGGGAACCATACCTGCTCTTTCAGAGAGGATTTAAAGCTGATCAGCATCGCTGGAT 3960  
Qy 2707 GCACCGCTGATTAACGTCGTTGGTTCAGGTTATCAGGTGCGTTTCTGCGCTGCGCG 2766  
Db 3961 AGCGCGCTGTGGTCCATCACTGCGCGGACTACCAAGTTGGTTTCCCCCAACGCGCTCG 4020  
Qy 2767 CAATCTGGAGCGGATTTATGTTCCATCGCATCAGCGGATTTTCCCGCGCATTTTACGC 2826  
Db 4021 CCATGTGAACAGTGGCTACTACTGCGTGACCTCCCGGCAATTCGCGCGGATACTCGGCA 4080  
Qy 2827 GCGATGGGTGACGATCTGTGACAAACACAGCGCTACACAGGTAAACCCACGCGAGT 2886  
Db 4081 ACAGTTTGAACAACATTTATGGCTGATACCGCGGTTTCAGCGCTTCATGCTGAATCGGT 4140  
Qy 2887 GACGCTGGCGGATGGCGGTGAATCTGCTGCGAAGTGGTGAATGATGGTGGCGGCTGCA 2946  
Db 4141 CCAGTTAGCGGATGGCGGATTTATCCAGTACAGTGAACGAGCGGAGGTTTACAC 4200  
Qy 2947 GCGGACGCACATCTCGAGTGGTTATCAGGTGTTTCTTGACAGAGTGGCAGCTGGC 3006  
Db 4201 GCTCGATTCTGCACTACGCTAGGATTCAGGCAATTTATCGGTGAGGAGTGGCACTGAG 4260  
Qy 3007 GCAGCGCAACGCGCTGCAGCAGCGGATCTGATGATGCCACGTCGATCAGCAACGCGG 3066  
Db 4261 CGGCGCGCATGGTTTATGTCACCGATTTATGATGATGGACGCTGATCAGCAAAATGG 4320  
Qy 3067 TTATCGTTTTGCTACACGTCGCGCTCAGCGCGGATCGGCTATTGATTTGAAGATACCCA 3126  
Db 4321 CTACCGCTTTGTTTATACCTTGGCTTTCGCAACCGCACTGCTGATCGAAGACACACA 4380  
Qy 3127 TTACGTTAACGACCGCGCTGGCGGAGAACACCGCTCTGTGACGACATCGCGGACTATGC 3186  
Db 4381 CTACATTGACAGGCTAAATCTTCAGGCGGAAACGCGCGCGCTCAGAAATTCGCGGATTCG 4440  
Qy 3187 CAATCAGCAGGCTGACGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3246  
Db 4441 TGGCGACAGGTTGGCGGTTACAGACGTTTCTGCGGAGAACAGGGTGCTATGGCCAT 4500  
Qy 3247 TACCTCTGAGCGGCAACATCGATCGATTCTGGCAACAGCAGCGCGGCAAGCGTGCAGCGG 3306  
Db 4501 TAGCTTAACGGCGGATAATCGTCAGTTTGGCAACAGCAAC--CGCAAGCTGTAGCGG 4557  
Qy 3307 CTTGCGCGCGGCTGTTTATCGCACACCGGTTTACTCTTGGCGCTCGCGCTGCGCTGCGCT 3366  
Db 4558 ATTACGCGCGGCGCTGTTTATCCGACAAACGCGCTACTCCCTACCGCTCGCGTGGCGCT 4617  
Qy 3367 AGCGGATGGTAGCAGCGCTGTGCGCACCGATGCCCTCAGCTCAGCGCAACATATCGA 3426  
Db 4618 GCGCGATCGCTCAGCGCGCTGGATGTGTTTACCTCTTCTCTGTTTACACAGCAGGATTCG 4677

Qy 3427 AGCTTTGCCCCCTCAGCAGTGGCGGAAACAGCGAATTTTTCCGTCTGCTAAACCGCATGCT 3486  
Db 4678 TCACITTTGCCAGCAACGTTGGCAGCAACAGGGGTTTTTCCGCACTGCTGAATCGCATGTT 4737  
Qy 3487 GTTTTTGGCCCGTAAGCCGACGCGCTGGCGGCTGATGACACGTTTTTACCGGCTCGA 3546  
Db 4738 GTTTTTAGCCGACCGCGGCTCAGCTGGGCTGTGATGACAGCTTTCTATGGCTTACC 4797  
Qy 3547 TGCCGGGTTAAATTAGCCGCTTTTACGCCGGCAACTGCGCCCTGCGGATATAAACCGCGAT 3606  
Db 4798 CGAGGATTTGATTTGCCGCTTTTATGCGGGAACCTCACCGTGACCGATCGGCTACGCT 4857  
Qy 3607 TCTGTGGCGCAAGCGCGCTGCCCATTCGGTGAAAGCGCTGCGCGCTGTGTGAATTCGT 3666  
Db 4858 TCTGAGCGCAAGCGCGCTTCCGCTTTTCGCGCATTTGACAGGCAATTTATGACGACTCA 4917  
Qy 3667 CGAACCGAGGAAGAAATGAACGCACTTATGTGATTTGGCGCAGGCTTTGCGGCGCTG 3726  
Db 4918 TCGTTGAAGAGCGACTACATGAACCAACTACGGTAATTTGGTGGCGGCTTTTGGTGGCTG 4977  
Qy 3727 GCGCTGGCGATTTCGCTGCAAGCGCGGCGCATACCAACCACTTACTCGAGCAGCGCGAC 3786  
Db 4978 GCACTGGCAATTCGTTTACAGCGCGCAGGTATTCCTGTTTCTGCTGCTTGGCAGCGGAC 5037  
Qy 3787 AAACCGCGCGAGCGCGCTTATGTTTGAAGACAGTGGCTTTTACCTTCGATGCGGACCC 3846  
Db 5038 AAGCGGGTGGCGGCTTATGTTTATCAGGAGCAGGCTTTTACTTTTATGATGAGCGCCCT 5097  
Qy 3847 ACCTGTATCACCAGTCCCGATCGAGCGGCTTCTATCGCTGTGCTGGGAAGATGGC 3906  
Db 5098 ACCGTTATCACCAGTCCCGATCGAGCGGCTTGAAGAACTGTTTGGCTTGGCGGCTTAAACGCTT 5157  
Qy 3907 ACCGATTACGCTGAGCTGATGCGGTAAACGCTTCTATCGCTGTGCTGGGAAGATGGC 3966  
Db 5158 AAGGATTAGCTGAGCTGTGGCGTACGCGCTTATCGCTGTGCTGGGAGTCCGCG 5217  
Qy 3967 AAACAGCTTGATTACGCAATTAATCAGCGCTGTGCTGGAGCAGAGATCGCCACGTTCAAT 4026  
Db 5218 AAGGCTCTCAATTACGATAACGACAGCGCCAGTTTGAAGCGCAGATACAGCAGTTTAAAT 5277  
Qy 4027 CCGCAAGATGTAGAAGCTATCGTCAATTTCTTGGCTATTCACGTGAATTTTACAGAG 4086  
Db 5278 CCGCGGATGTGGCGGTTTATCAGCGTTCTTGAATTTCTGCTGCTGCTTCAATGAG 5337  
Qy 4087 GGTATCTGAACACTCGGACCGTGGCTTCTGCAAGTGGCTGACATGTGCGCGCTCGCG 4146  
Db 5338 GGTATCTGAAGCTCGGCACTGTGCTTTTATCGTTCAAGACATGCTTCGGGCGCGG 5397  
Qy 4147 CCGCAGTTGGGACGCTGCAAGCATGGCGCAGCGCTTACAGCATGGTGGCGAAATTTAT 4206  
Db 5398 CCCCAGTTGGCAAGCTGCAGGCAATGGCGCAGCGTTTACAGTAAAGTTTCCGCGCTACAT 5457  
Qy 4207 CAGGAGCATCATCTGGCTCAGGCTTCTTCTCACTCATTTGCTGGTGGCGGCTATCCT 4266  
Db 5458 GAGGATGAGACTTCTGGCAGCGGTTTCTTCTTCACTCGCTTCTAGTGGGGGGAATCCG 5517  
Qy 4267 TTTGCAACGTCATCATCTATACCTTAAATTCATGCGCTGAGCGCTGAATGGGCGCTG 4326  
Db 5518 TTTGCAACCTCTCTCATTTTATACGCTGATTCACGCTTGAAGCGGAAATGGGCGCTG 5577  
Qy 4327 TTTCCGCGCGCGGCAACCGCGCTGTGTGAGGCGATCGCGCATCTGTCGAGGACTTG 4386  
Db 5578 TTTTCAACGCGGTGAAGAACCGCTGCTGCTTCTTCTTCACTCGCTTCTAGTGGGGGGAATCCG 5637  
Qy 4387 GCGCGGAGCTGTTTACTGATGCGGAGTGAAGCAGCTGGAACCAACCGCGCATCCGAT 4446  
Db 5638 GCGCGGAGCTGCTGCTTAAACCGCGGCTCAGTCAATGGAACCGTTGGGGAACAAGATT 5697  
Qy 4447 AGCGGCTTCAGTTAGAGGCGGACGACGCTTTCGATGCGCGCTGTGCGCTTCAATGCC 4506  
Db 5698 CAGGCGCTGCACTTGAAGACCGGACGCGTTTGAACCTGCGCGTGGCGTGGACGCT 5757



QY	4507	GAGTGGTGATACCTACGACAAACCTGCTTCGCCACCATCCGCTGGCGAATGAAACCTGGC	4566
Db	5758	GATGTTGTATACATACCTATCGCGCATCTGCTGTCTCAGCATCCCGCACCGCTAAGCAGGGC	5817
QY	4567	ACATCCGCTGAAGCGTAAAGCGCATGAGCAACTCGCTGTTTGTACTCTATTTTGGCCCTGAAT	4626
Db	5818	AAAAAACTGCAATCCTCAAGCGTATGAGTAACCTCACTGTTTGTACTCTATTTTGGTCTCAAC	5877
QY	4627	CAGCCGCATGAACAGCTCGCGCACCAACACACGCTCTGTTTGGCCCGCGTTATCGTAGTTG	4686
Db	5878	CATCATCAGCATCAAACCTCGGCCATCATACCGTCTGTTTGGGCGACCGTACCGTGAACCTG	5937
QY	4687	ATCGATGAGATTTTCAACAGCAGCCAGCTGGCAGACGATTTTTCATTTTACCTTGCACCGG	4746
Db	5938	ATTCAAGAAATTTTAAACCATGATGCTGGCTGAGGATTTTTCGCTTTATTTTACACGCA	5997
QY	4747	CCCTGCAGCAGCGATCCGTCGCTGGCAGCCGCCCGCGCTGGCGCAGCTTTTATGTGTTAGCG	4806
Db	5998	CCTTGTGTACGGATCCGTCACTGGCACCGGAAGGTTGGCGAGCTATTATGTGCTGGCG	6057
QY	4807	CCGGTGGCGGATCTCGGCACCGCTGACATCGACTGGCAACAGGAAGGACCGCGCTTGGCG	4866
Db	6058	CCGTGTTCCACACTTAGGCACCGGCGAACTCTCGACTGGCGCGTAGAAGGACCCGACCTGCGC	6117
QY	4867	GATCGAATTTTGTCTTCTGAGAGCAGCACTACATCCGGGATTAGCTCAGCAATTTAGTG	4926
Db	6118	GATCGTATTTTGACTACCTTGAGCAACAATTAATGCTTGGCTTGGGAAGCCAGTTGGTG	6177
QY	4927	ACACACAGAAATGTTTACCGCGTTTGATTTTCGGGACACAGCTGCAATCGCCATACGGCTCG	4986
Db	6178	ACGCACCGTATGTTTACCGCGTTTCGATTTCCGCGACGAGCTCAATGCTGCGCAAGGTTTCG	6237
QY	4987	GCCTTTTCGCTGAGCGGATTTTGACGCAAGCGCTGTTTCGCCCGCATTAACCCGAT	5046
Db	6238	GCCTTCTGGTTGAAACTATTTGACCCAGAGCGCTGTTTCGACCAATAAACCGCGAT	6297
QY	5047	GCCGATATACGAATCTCTATCTGTTGGTGGCGGTACGCAATCCAGCGCGGCGCTGCC	5106
Db	6298	AAGCACATTGATAATCTTTATCTGTTGGCGCAGGCAACCCCATCTCGGCGCGGCGATTTCCC	6357
QY	5107	GGCGTATCGGTTTCGGCCAAAGGCCACCGCCAGGCTGATGCTGGAGAGATCCGCCGGAATGA	5166
Db	6358	GGCGTAATCGCTTCGGCGAAGGCGACGGCAGGCTTAATGCTGGAGGACCTGATTTGACGA	6417
QY	5167	ATCGACAGCCTTTACTTTGAGCAAGTAAACGCAAACCAATGGCGGTGGGCTCGAAGAGTTTCG	5226
Db	6418	AT---ACGTCAATTACTGAATCATGSCCGTCGAAACCAATGGCGTTGGCTCGAAAGGCTTTG	6474
QY	5227	CCACCGCGCCAGCTGTTTGTATGCACCGGCGCGCAGCAGCTGATGCTGATGCTGATGCGT	5286
Db	6475	CGACTGATCGACGCTTTTCGACGCGCAAAACCCGCTCGCAGCGTCTGATGCTTTTACGCA	6534
QY	5287	GGTGTCTGCTACTGCGATGATGTGATTGATGGGCAGAAACGCTGGGCGAAGCGCGCACGACG	5346
Db	6535	GGTGGCCGCACTGCGACGACGTCATTGACGATCAAACACTGGGCTTTTCATGCCGACCGC	6594
QY	5347	ATGCCGTCGAAGACCGCAGCGACGATATGACGATCTGCAAAATGGAACCCGCCGCGCCT	5406
Db	6595	CCTCTTCCGAGATGCCCTGAGCAGCGCCTCGACGAGCTTGAATGAAACCGCGTCAGCGCT	6654
QY	5407	ACAGCGCGCGCACATGATGAACCGCGGTTTATGGGCGTTTTCAGGAAGTGGCGATCATTC	5466
Db	6655	ACGCCGGTTGCAAAATGACAGAGCCCGCTTTTGGCGGTTTTCAGGAGTTCGCGATGGCGC	6714
QY	5467	ACCAAGCTGCCGCAACAACTGCGGTTTGCATCATCTGGAAGGCTTCGCTATGGATGCACGCA	5526
Db	6715	ATGATATCGCTCCCGCCTAACGCGTTTCGACCATCTGGAAGGTTTTCGCCATGATGTGCGCG	6774
QY	5527	ACGAAACATTACCGGAGCTTCGATCAGACGCTGCGTTACTGCTATCAGTTCAGTTCGCGGCGTGG	5586
Db	6775	AAACGCGCTACTCTCAGCATGACGATGACGCTGCGCTTATTTGCTATCAGTTCGCGCGGTTG	6834
QY	5587	TCGGTTTGTATGATGGCGCGGCTAATGGGCGTGGCGCAGACGAGCGGTGCTCGATCACGCT	5646

Db	6835	TGGCCCTGATGATGGCCAAATTAATGGCGTTTCGGATACGCCACGCTCGATCGCGCT	6894
Qy	5847	GCATTTTAGGACTGGCGTTCCAGCTCACTAACATTCGGCGCGACATGTAGAAAGATGCCG	5706
Db	6895	CGCATTCGGCGTGGCTTTCAGTTGACCAACATTCGCGGTGATATTGTTCGACGATGCTC	6954
Qy	5707	AAAATGTCCTGCTATCTGCGCAATCTCGGCTCGATCAGCGGGATTAACGCCGATA	5766
Db	6955	AGTGGCGCGCTGTATCTCGCTGAAGCTGGCTTGGAAAGGAAGGACTGACAAAGCGA	7014
Qy	5767	CGCTGACTGCACCGCAACATCTGTGCAGCGCTCGCCTCACTGGCAGCGCGTTTGTGGCGG	5826
Db	7015	ATTATGCTGGCGCAGAAAACGGCAGGCGCTTAAGCGGTATCGCGGCGACTGTACGGG	7074
Qy	5827	AGCGGAAACCTTATTATCACTCGCGCGGATCCGGTTTACGGGTTTACCGCTCGGCTCGG	5886
Db	7075	AAGCGGAACCTATTATGCTATCATATGATCGCGGTCTGGCACAATTACCCTTACGCTCGG	7134
Qy	5887	CGTGGGCCATCGCTACGCGCTCGCGCGTTTATCGCAAAATTTGGCGTCAAAAGTTTCAACG	5946
Db	7135	CCTGGGCATCGCAGACGACGAGCGGTGTACCGTAAATTTGGCGTCAAAAGTTTGAACAGG	7194
Qy	5947	CCGCTGTGCACGCTCGGATTTACGCGAGCGCACAGTAAAGGTGAAAACTCGCGCTGC	6006
Db	7195	CCGTTAAGCAGGCTCGGATCATCGCCAGTCCACGTCACCGCGCGAAAAATTAACGCTTT	7254
Qy	6007	TGGTGAAGGGCAGGTTTCGGCATCACTTCGGGTGTCTCGTCTCTGAACCGCGTCCGG	6066
Db	7255	TGCTGACGGGCATCCGGTCAGCGAGTTACTTCCCGGATGAAGAAGTATCCACCCGTCCTG	7314
Qy	6067	CTGCTCTGTGGCAGCGTCTCTGTTGATTTTACGTCCTGTGACGCTGGCGCAGCGTGGCTTG	6126
Db	7315	CTCATCTCTGGCAGCGCCC-----GATCTAGCGCGCATGCTTCTCTCAGCGTGGCGTG	7368
Qy	6127	CAGCTTATTACGGGTGGCGGTAGAGAAACCAACACACGACGCGCTTTCACGCCCGCG	6186
Db	7369	AAGTTTAGATAACGGTCGGCGGTACAGAAAACCAAGAGNACGCGAGCCCTCTTTTCCCT	7428
Qy	6187	CACCGCATGATGCGGTGGCGCATGTATTAAGCGCTTAAGATAGGCTTTGGCGCGGAT	6246
Db	7429	TACAGCATGATGATACGGTGGGCCATGTATTAACCGTTTCAAGTAGCTTTGGCGCGTAT	7488
Qy	6247	ATAGCGAAACGGCAGCGTTGATGCAACAGCGCATCTGTGCACCATGAAGTAGAGCGGCC	6306
Db	7489	GTAGCGAAACGGCAGCGCTGGTGTACAGTCCGCTGTGGACCAATAAAATACAGTAAACC	7548
Qy	6307	GTAGTCTGTCACTCCGCGACCAATTCACATGCGAGCGGCCACATGCTTGCACACCGCAT	6366
Db	7549	ATAAGCGGTATGCTGCAACCATTCATCGAGCGGCCAGATTCTGTACTGTCGCGAAGTA	7608
Qy	6367	AATCAGCACAATCGCCAGTACCGCAAAACACACCGCATAAAGATCGTTGTAGCTCAAACTT	6426
Db	7609	AATCAGGSCAATCGACACAAATGGCGAATACCAACCGCATAGAGATCGTTTAACTTCAATGC	7668
Qy	6427	ACCGCTGTGCGTTTCATGGTTCGACAGATGCGACGCCCATCCCAACCGTGCATGAGTA	6486
Db	7669	GCCTTTACGCGGTTTCATGATGTGAAGATGCGACCCCAACCCAGCGCGTGCATGATGTA	7728
Qy	6487	TTTATGCAACAGCGCGCTACGATTTCCATCACACCAACCGTTCGCCAACAGATAAGCAC	6546
Db	7729	TTTATGTCGAGTGCAGCAACCACTTCCATGCGCACCAACCGGTGACAAAACGATCAGGGC	7788
Qy	6547	GTTCCATTAACGAGCATGTGTTGTTCCA	6574
Db	7789	ATTCCAAATCCACAACATAATTTCTCA	7816

RESULT 5  
ADQ48673  
ID ADQ48673 standard; DNA; 8609 BP.  
XX AC ADQ48673;

Qy





QY 1810 ATTGACGGAGCGGCGGATTAATCATAGTGCCTCTGCGCGTGCACAAATCAGCCAGA 1869  
DB |||||  
QY 3062 GTTATAGCCCCGGGTAAGAAATGCAATCTGTTTCTCCACTGGCAAAATCAGCCAGT 3121  
DB |||||  
QY 1870 TGGTGGCGGCTTTGATTTTCCAGCTGACAACTGCCCGCTGCTATCACGCGGTTGGGC 1929  
DB |||||  
QY 3122 TGAATCCCGAATGATTTTCCCGAAGCGCTGCCAGACTGCTTTTATGGGTTGGAC 3181  
DB |||||  
QY 1930 CACTCCGCG- - - CCGCGGTTTCTCTGGCGGCTCCATGCGCCCTGGCCAGCGCTGCGTC 1986  
DB |||||  
QY 3182 CGTTACGGCAACCCAGGGAGCGCGGGGTCACTCAATCTTTATTTCCGTCGCCGACA 3241  
DB |||||  
QY 1987 AGCGGTGTTTATGCTGCTGGTACGCTGCGTCAAGGCCATCGCTTCGCGTGTCTGTC 2046  
DB |||||  
QY 3242 AACCCTGATTTTGGCTCGCTGGGACCCCTGCAGGGACATGTTATGGCCCTGTTTCAAGA 3301  
DB |||||  
QY 2047 ATCTGGCGCAGGCTGCGCCAGCTGCGGCTATCGCTGTGATCGCCCATTTGTGGGGAT 2106  
DB |||||  
QY 3302 CCATCGCCAAAGCTGCGAAGGTTGATGCGCAGTTACTGTTGGCACACTGTGGCGGC 3361  
DB |||||  
QY 2107 TAAAGCCGAAACAGACGCAATACGCTGAGTGGAGCTGCTGGCGCGCTGGGTGACGGATTTG 2166  
DB |||||  
QY 3362 TCTCAGCCACGACGAGGAGGTGAATGCGCCGGGCGGGGACATTCAGGTTGTGGAATTTG 3421  
DB |||||  
QY 2167 TCGATCAGCGCGCAGCCCTACAGCAGCGCAGCTGTTTATCACTCATGCCGGGTTAAACA 2226  
DB |||||  
QY 3422 CGGATCAATCCGAGACATTTTACAGGCACAGTTGACAAATCAACATGTGGGATGAATA 3481  
DB |||||  
QY 2227 GCGCGCTGGAAGCATTGGAATGCGGTAACGCGATGCTGGCGTCCGATTTGCTTTGATC 2286  
DB |||||  
QY 3482 CGGTACTGAGCGCTATTGCTTCCGCAACACCGCTACTGGCGTGGCGCTGATTTGATC 3541  
DB |||||  
QY 2287 AGCCCGGCTGGCGCGCGGATGAGTGGCATGACGTTGGTGGCGCGCGCATCACGCTTTA 2346  
DB |||||  
QY 3542 AACCTGGCGTGGCATCACGAATGTTTATCATGGCATCGGCAAGCGTGGCTCTCGGTTTA 3601  
DB |||||  
QY 2347 GCGGTGTCATCACTGAGGAGCAGCATCTGCAACAGCTGCTGACCGACAGTCTGTACGCGC 2406  
DB |||||  
QY 3602 CTACAGCCATGCGCTGGCGGCGCAGATTGATGCTGCTGCTGCTACTAACCCGATTACCGC 3661  
DB |||||  
QY 2407 TAGCGATGTACGCAATTCAGGCGCAGCTGACGCGCGCAGCGGTTGCCAGTGGCGCG 2466  
DB |||||  
QY 3662 AGCGTATGACAAAATTCAGGCGCATTTGGCTGCGCAGCGGCGCACACAGCGCGCGCG 3721  
DB |||||  
QY 2467 ACATGCTGAGCAGGCGCTGTGCCAGCAGTGTGCTGGCGGAGGCGACTGATGCG 2526  
DB |||||  
QY 3722 ATATTGTTGAACAGGCGATGCG- GACCTGTACGCCAGTACTCATGTGGGAGGATTAIGCA 3780  
DB |||||  
QY 2527 CAGGCAATACGATGTGATTTTGGTGTGCTGCTGAGCTGGCGAATGGCTTGATTCGCTGCG 2586  
DB |||||  
QY 3781 ACCGCACTATGATCTCATTTCTGCTGCTGGCTCTGGCTAATGGCTTATCGGCTCG 3840  
DB |||||  
QY 2587 TCTGCGTCAATTCAGCCCAACATGAAATGCTGTTGCTGGAGAGCGATGCGCATCCGCG 2646  
DB |||||  
QY 3841 GCTTCAGCAACAGCATCCGATATGCGGATCTTGCTTTATGAGCGGCTCTGAGCGCGG 3900  
DB |||||  
QY 2647 AGGCAATCACTATGCTGCTTTCATCACGCGATCTCAGCGCGCAACACTTCGCTGGCT 2706  
DB |||||  
QY 3901 AGGGAACCACTGCTGCTTTCACGAAGAGGATTTAAACGCTGAATCAGCATGCTTGGAT 3960  
DB |||||  
QY 2707 GCAACCGCTGATTTACCGTGTGGTTCAGGTTATCAGGTGCGTTTTCCTGCGTGGCGCG 2766  
DB |||||  
QY 3961 AGCGCGCTTGTGTCATCACTGCGCGGACTACGAGTTTCGTTTCCCCCAACCGCGTCG 4020  
DB |||||  
QY 2767 CAATCTGGACGGGATTAATTTGTCATCGCATCAGGGGATTTTCCGCGCATTTTACGC 2826  
DB |||||  
QY 4021 CCATGTGAACAGTGGCTACTACTGCTGACCTCCCGCATTTTCGCGGGATCTCCGGCA 4080  
DB |||||  
QY 2827 GCGCATGGGTGACGATCTGTGGACAAACACAGCGCTACACAGAGTAAACCCACGAGGT 2886  
DB |||||  
QY 4081 ACAGTTTGGACAACTTTATGGCTGTCATACCGGGTTTCAGCGGTTTCATGCTGAATCGGT 4140  
DB |||||

QY 2887 GAOGCTGGCGATGGCGCGTGAACCTTGTCTGCGCAAGTGGTGAATGATGCTGCGCGCTGCA 2946  
DB |||||  
QY 4141 CCAGTTAGCGGATGGCGGATTAATCATGCGAGTACAGTACGATCGACGAGCGGTTACAC 4200  
DB |||||  
QY 2947 GCGACGCCCAATCTGAGCTGGGTTATCAGGTGTTTCTTGGACAAGAGTGGCAGCTGGC 3006  
DB |||||  
QY 4201 GCCTGATTTCTGCATACGCGTAGGATTCAGGCAATTTATCGGTGAGGAGTGGCAACTGAG 4260  
DB |||||  
QY 3007 GCAGCCGACGGCTGCGAGCAGCGCTCTGATGATGATGCCACCGTCCATCAGCAAGCGGG 3066  
DB |||||  
QY 4261 CGCGCGCATGGTTTATCGTCACCGATTAATGATGATGCGACGCTGCTGATCGAAGACACACA 4320  
DB |||||  
QY 3067 TTATCGTTTTGTCTACACGCTGCGCTCAGCGCGGATCGGCTATTTGATGAAGATACCCA 3126  
DB |||||  
QY 4321 CTACCGCTTTGTTTATACCTGCGCTTCCGCAACCGCACTGCTGATCGAAGACACACA 4380  
DB |||||  
QY 3127 TTAGCTTAAACAGCCCGGCTGGCGGAGAACACCGCTGCTGACGACATCGCGACTATGC 3186  
DB |||||  
QY 4381 CTACATTTGACAAGGCTAATCTTACGGCCGAAACGGGCGCTCAGAAACATTCGCGATTATGC 4440  
DB |||||  
QY 3187 CAATCAGCAAGGCTGGACGCTGAGTACGCTGCTGCTGAAAGAGCACGCGCATATTACCGAT 3246  
DB |||||  
QY 4441 TCGCGGACAGGTTGGCGGTTACAGAGCTTGTCTGCGGGAAGAACAGGTTGCAATTTGCCAT 4500  
DB |||||  
QY 3247 TACCTGAGCGGCAACATTCGATTCGTCGAAACAGAGCGCGGCCAAGGTCGACGCG 3306  
DB |||||  
QY 4501 TACGTTAACGGCGGATTAATCGTCACTGTTTGGCAACAGCAAC- - - CGCAAGCTGTAGCGG 4557  
DB |||||  
QY 3307 CTTGCGCGCGGCTGTTTTCATGCGCACACCGGTTACTCTTTCGCTGCGCGTGGCGCT 3366  
DB |||||  
QY 4558 ATTACGCGCGGCTGTTTTCATCGCAACCGGCTACTTCCCTACCGCTCGCGTGGCGCT 4617  
DB |||||  
QY 3367 AGCGGAGTTGTAGCAGCGCTGTTTGGCCCAACCGATGCGCTCACGCTCAGCCAAACATATCGA 3426  
DB |||||  
QY 4618 GCGCGATCGTCTCAGCGCGCTGATGTTTACCTTCTCTCTGTTTACCAGACGATTCG 4677  
DB |||||  
QY 3427 ACCTTTGCGCGTACAGAGTGGCGGCAACAGCGATTTTTCGCTGCTGCTAAACCGCATGCT 3486  
DB |||||  
QY 4678 TCACTTTGCCAGCAACGTTGGCGCAACAGGGGTTTTTCCGCACTGCTGAATCCGATGTT 4737  
DB |||||  
QY 3487 GTTTTGGCCGTAAGCCGACGAGCGCTGCGGTGATGCAAGCTTTTACCAGCTGCA 3546  
DB |||||  
QY 4738 GTTTTGGCCGCAACCGCGCGAGTCAAGCTGCGGTGATGAGGCTTTTCTATGCTTACC 4797  
DB |||||  
QY 3547 TCGCGGTTAATTAGCGCTTTTACGCGGCAACTGCGCTGCGCGGATAAAAACCGGAT 3606  
DB |||||  
QY 4798 CGAGGTTTGTGCTGCGCTTTTATGCGGAAAACTCACCGTGAACCGATCGGCTACGCA 4857  
DB |||||  
QY 3607 TCTGTGCGCAAGCGCGGTCCTTCCATCGGTGAAGCGCTGCGCGCTGTTGAATTCGT 3666  
DB |||||  
QY 4858 TCTGAGCGGCAAGCGCGCTTCCCGTTTTCGCGCATTTGAGGCAATTAAGACACTCA 4917  
DB |||||  
QY 3667 CGAACACGGAAGAAAAATGAACGCACTTATGATTTGGCGCAGGCTTTGGCGGCTG 3726  
DB |||||  
QY 4918 TCGTTGAAGAGGACTACATGAACCACTACGTTAATTTGGTGGGCTTTTGGTGGCTG 4977  
DB |||||  
QY 3727 GCGTGGCGATTCGCTTCAAGCGCGGCAATCAACACACTTACTCGAGAGGCGGAC 3786  
DB |||||  
QY 4978 GCATGGCAATTCGTTTACAGCGCGAGTATTCCTGTTTGTGCTGTGAGCAGCGGAC 5037  
DB |||||  
QY 3787 AAACCGGCGGACGCGCTATGTTTTCAGGACAGTGGCTTTTACCTTCGATCCCGGACCC 3846  
DB |||||  
QY 5038 AAGCGGGTGGCGGCTTATGTTTATCAGGACAGGCTTTTACTTTTGTGAGCGGCT 5097  
DB |||||  
QY 3847 ACCTGTGATCACCGATCCAGCGCCATCGAAGATTTGTTTACGCTGGCAGGAAAAATCGCTC 3906  
DB |||||  
QY 5098 ACCGTTATCACCGATCCAGCGGATTAAGAACTGTTTGTCTGCGCGGTAACAGCTT 5157  
DB |||||  
QY 3907 AGCGATTACGTCGAGTGAATGCGGTAAGCGCTTCTATCGCTGCTGGGAAAGATGCG 3966  
DB |||||  
QY 5158 AAGGATTACGTCGAGCTGTGCGGCTACGCGGTTTTATCGGCTGTGCTGGGATCCGCG 5217  
DB |||||  
QY 3967 AAACAGCTTGATTACGACAAATAATCAGCGCGCTGCTGGAGCAGCAGATCGCCACGTTCAAT 4026  
DB |||||









Db 8627 GAACGGCATGTTGAAGCTGTTTACCGATCTGGCGGGGAGATCGAACTCAACGCCGGGT 8686  
Qy 4416 GAGCAGCTGGAAACACAGCGGCAATCGCATTTAGCGGCTTCAGTTAGAGGGCGGACGACG 4475  
Db 8687 CGAAGAGCTGGTGGCGGATTAACCGGTAAGCCAGGTCGGCTGGCGGATGGTCGGAT 8746  
Qy 4476 TTTCGATGCCCGCTGTGGCTTCAATTCGCGACGTGTGCANATCCTACGCAAACTGCT 4535  
Db 8747 CTTTGACACCGACGCGTAGCTCGTGAACGCTGACGTGTGAACCTATATAAAGCTGCT 8806  
Qy 4536 TGGCACCATCCGCTGGCAATGAACCTGCGACATCGCTGAAGCTGAAGCGTAAGCGCATGACAA 4595  
Db 8807 CGGCAACCATCCGGTGGGCGAAGCGGGCGGACGCTGGAGCGCAAGAGCATGAGCAA 8866  
Qy 4596 CTCGCTGTTTGTACTTATTTGGCTGAATCAGCCGATGAACAGCTCGGCGACACAC 4655  
Db 8867 CTCGCTGTTTGTCTTACTTCGGCTGNAACAGCTCATTTCCAGCTGGCGCACCATAC 8926  
Qy 4656 CBTCTGTTTGGCCCGCTTATCGTGAAGTTGATCGATGAGATTTTCAACAGCAGCAGCT 4715  
Db 8927 CATCTGTTTGGTCCCGCTACCGGAGCTGATCGACGAGATCTTTACCGGCGCGCT 8986  
Qy 4716 GGCAGACGATTTTCACTTACTCGACGCGCTCGACGAGATCCGTGCTGCGACC 4775  
Db 8987 GCGGATGATTTCTCGCTCTACTCGCTACCTGCGCTGCGTGACCGATCCCTGCTCGGCC 9046  
Qy 4776 GCCGCTGCGGCGAGCTTTATGTGTTAGCGCGGTGCGCATCTCGGCGACCGCTGACAT 4835  
Db 9047 TCCCGCTGCGCAGCTTCTAGTGTCTGCTGCGCCCGGTGCGCATCTTGGCAACGCCGCT 9106  
Qy 4836 CGACTGGCAACAGGAAGACCGGCTTGGCGCATCGAATTTTGTCTTATCTGGAGAGCA 4895  
Db 9107 GBACTGGGCGCAGGAGGGGCCGAAGCTGCGCGACCGCATCTTTGACTACTCTTGAAGAGCG 9166  
Qy 4896 CTACATGCGGATTTACGTACGATTTAGTGACACAGAAATGTTTACGCCGCTTTGATTT 4955  
Db 9167 CTATATGCGCGGCTGCGTAGCGAGCTGGTGACCCAGCGGATCTTTTACCGCGGAGACTT 9226  
Qy 4956 TCGGCACACGCTGCATGCCATCACGGCTCGGCGTTTTTCGCTGGAGCCGATTTGACGCA 5015  
Db 9227 CCACGACACGCTGGATGCCATCTGGATCGGCTTCTCCATCGAGCGCTGCTGACCCA 9286  
Qy 5016 AAGCGCTGGTTCGCGCCCGCATACCGCATGCGGATTCAGCAATCTCTATCTGGTGG 5075  
Db 9287 AAGCGCTGGTTCGCGCCCGCATACCGCATGCGGATTCAGCAATCTCTATCTGGTGG 9346  
Qy 5076 TGCCTGTACGATCCAGGCGCGGCTGCGCGGTGATCGGTTTCGCGCAAGCCACCGC 5135  
Db 9347 CGCAGTACTCACCTTGGGGGGGCAATCTGCGGTAGTGGCTCGGCGAAGCCACCGC 9406  
Qy 5136 CAGGCTGATGCTGGAGGATCGCGCCGAATGAATCGACAGCTTTTACTTGAGCAAGTAAG 5195  
Db 9407 CAGCCTGATGATTGAGGATC---TGCAATGAGCAACCGCGCTGCTTCAACCGCCACG 9463  
Qy 5196 CAAACATGGGCTGGCTGGAAGATTTGCGCAACCGCGCCGCAAGCTGTTGATGACCG 5255  
Db 9464 CAGACATGGCCAACCGCTCGAAAGTTTTTGCACCGCTGCGAAGCTGTTCGACCGCGCC 9523  
Qy 5256 AGCGCGGACGACGCTGATGCTGATGCTGTGCTGCTCACTGCGATGATGATTGAT 5315  
Db 9524 ACCCGCTGATGCTGATGCTTCTTACCTGCTGCGGCACTGCGATGACGCTATTGAC 9583  
Qy 5316 GGGCAACGCTGGGCGAAGCGCGCACGACGATGCGCTGAAAGCGCGCAGGCGAGTATG 5375  
Db 9584 GACCAGACCCAGGCTTCGCGACGAGGCGCGCGGCGGAGGAGGAGGCCACCCAGCGCTG 9643  
Qy 5376 CAGCATCTCAATTAACCCCGCGGCTACAGCGGCGCGACATGATGACCGCG 5435  
Db 9644 GCCCGCTGCGCACGCTGACCTTGGCGGCTTTGAAGGGGCGGAGATGAGGATCCGGCC 9703  
Qy 5436 TTTAGGCGCTTTCAGGAAGTGGCGATCATTCACGCTGCGGCAACAACTGGCGTTTGTAT 5495  
Db 9704 TTCGCTGCCCTTTCAGGAGGTGGGCTGACCCAGGTTATTAGCGCCCGCATGGCGCTCAT 9763

Qy 5496 CATCTGAAGGCTTCGCTATGATGATGACGCAAGAAATTAACGCGAGCTTCGATGACAG 5555  
Db 9764 CACTCGACGGCTTTCGATGAGGAGTGGCTCAGACCGCTATGTCACTTTTGAAGATACG 9823  
Qy 5556 CTGCGTTACTGCTATACGCTCGGCGGTGCTGTTTGTATGAGGCGCGGTAAATGGGC 5615  
Db 9824 CTGCGCTACTGCTATACGCTGGCGGCTGCTGGGTCTGATGATGGCCAGGCTGATGGC 9883  
Qy 5616 GTGCGCAGCAAGCGCTGCTGATCAGCCCTGCGATTTAGGACTGGGCTTCACGCTCACT 5675  
Db 9884 GTGCGGATGAGCGGCTGCTGATCGCCCTGCGATCTGGGCTGGCTTCACGCTGACG 9943  
Qy 5676 AACATTCGCGCAGATTTGTAGAAAGATGCCAAATGCTGCTGCTATCTGCCGCAATCC 5735  
Db 9944 AATATCGCCCGGATATTAATGACGATCGGCTATTGACCGCTGCTATCTGCCCGCCGAG 10003  
Qy 5736 TGGCTCGATCAGGCGGATTAACGCGCGGATACGCTGATCGACCGCAACATCGTGCAGCG 5795  
Db 10004 TGGCTGAGGATGCCGCTGACCCCGAGAACTATGCGCGCGGAGAAATCGGCGCGC 10063  
Qy 5796 CTCGCTTCACTGGCAGCGGCTTTAGTGGCGAGCGGAACCTTATTAATCACTCGCGCGCA 5855  
Db 10064 CTGGCGGGTGGCGAGCGGCTTATTGATGCGCAGAGCGCTACTACATCTCTCCAG 10123  
Qy 5856 TCCGCTTACCGGCTTACCGCTGCGCTGCGGCTGGGCGCATCGCTACGGCTCGCGGCTT 5915  
Db 10124 GCCGCGCTACAGATGCGCGCGCTGCGCTGGGCGATCGCACCGCGCGCAGGCTC 10183  
Qy 5916 TATCGGAAATTTGGCGTCAAAAGTTTTCAGCACCGCGTGTGACGCTGGGATTCACGGCAG 5975  
Db 10184 TACCGGAGATTCGGTATTAAGGTAAAGCGCGGAGCGCGCTGGGATTCGCGCGCAG 10243  
Qy 5976 CGCACAGTAAGGTGAAAACTGCGCTGCTGTAAGGAGGCGAGGCTTTGGCGATCACT 6035  
Db 10244 CACACGCAAGGTGAAAAATTTGCCATGCTGATGGCGCACCGCGGCGAGCTTATCGG 10303  
Qy 6036 TCGGCTGCTGCTGCTGTAACCGGCTCGGCTGCTGTTGGCGAGCTTCCTGTTGATTT 6095  
Db 10304 GCGAAGAGCAGCGGCTGACCGCGCTCGCGCGCTTTTGGCAGCGCTTCGTTAGGCG 10363  
Qy 6096 TACGCTCGTACGCTGGCGCAGCGTGGCTTCAGCTTATTCAGCGGTGGCGCTAGAGGA 6155  
Db 10364 GCGCGCATGATCTTCAGCGAGATTCGCTGATGCTGGCGAGCTTCGCGGCGTAAATAA 10423  
Qy 6156 AACCAACGACACGCGCTTTCAGCGCGCGGACCGCATGATGATGCGGTGCGGCTGCT 6215  
Db 10424 AACCGAAGAGCAGCGCTTCCGCGCGCGCATCGCGCTGCTGCGAGCGCTTCGCGG 10483  
Qy 6216 ATAAGCGTTAAGATAGCTTTTGGCGGATATAGCGAAGCGCGCGGCTGATGACCA 6275  
Db 10484 AGAGCGCTTTCAGGTAGCCCCGCGGAGTCCAGTGAAGGCGCGCTGATGACCA 10543  
Qy 6276 GCGCATCTGACCAATGAAGTAGAGCGCGCTTACGCTCATTCGCGCAGCAATCCACT 6335  
Db 10544 GACCGCTGTCACAGAGAGTAGAGCGGCTAGACCGCTCATGCGCGAGCAATCCACT 10603  
Qy 6336 CGAGCGGCAATGCTTTCAGCACCGCAATTAATCAGCAATCGCGAGTACCGCAACA 6395  
Db 10604 GCGGCGCAACCGCGCGCTGCGCAATCAGCGGATAGCGGCGGCGGCAACA 10663  
Qy 6396 CCACCGCATTAAGATGCTTGGCTCAAACTTACCGCTGCGGTTTCATGCTGCGAGAT 6455  
Db 10664 CCACCGCAAGAGATTCGTTTACCTCAATAACCGCTTGGCGGCTTGGTGTGCTCAT 10723  
Qy 6456 GCCAGCGCATCCCAACCGCTGATGATGATTTATGCGACAGCGCGCTACGATTTCCA 6515  
Db 10724 GCCAGCGCATCCCAACCGCTGATGATGATGATGCGGTGGGTAACCGCGGATGCCCTCA 10783  
Qy 6516 TCACCAACGAGTTCGCAACAGATAA 6542  
Db 10784 TCGCAATAACGCTCAAGATGACGATTA 10810



RESULT 7  
ADP74122  
ID ADP74122 standard; DNA; 4624 BP.  
XX  
XX AC ADP74122;  
XX DT 26-AUG-2004 (first entry)  
XX DE Pantoea ananatis DNA fragment encoding CrtY, CrtI and crt8.  
XX KW ketocarotenoid; ketolase; carotenoid; hydroxylase; beta-cyclase;  
XX KW astaxanthin; adonirubin; canthaxanthin; adoxanthin; zeaxanthin;  
XX KW antioxidant; pigment; animal feed; trout; salmon; shrimp;  
XX KW food supplement; gene; ds.  
XX OS Pantoea ananatis.  
XX FH Key Location/Qualifiers  
XX FT CDS 128..1267  
XX FT /\*tag= a  
XX FT /product= "CrtY"  
XX FT 1288..2766  
XX FT /\*tag= b  
XX FT /product= "CrtI"  
XX FT 2802..3692  
XX FT /\*tag= c  
XX FT /product= "crt8"  
XX DE10253112-Al.  
XX PN  
XX XX  
XX PD 03-JUN-2004.  
XX PF 13-NOV-2002; 2002DE-01053112.  
XX PR 13-NOV-2002; 2002DE-01053112.  
XX PA (SUNG-) SUNGENE GMBH & CO KGAA.  
XX PI Sauer M, Flachmann R, Klebsattel M, Schopfer CR;  
XX WPI: 2004-489014/47.  
XX P-PSDB; ADP74123, ADP74124, ADP74125.  
XX PT Production of ketocarotenoids with low hydroxylated by-product content,  
XX PT for use e.g. in pigmenting feedstuffs, by culturing genetically modified  
XX PT organisms having modified ketolase activity.  
XX PS Example 2; SEQ ID NO 24; 101pp; German.  
XX  
CC This invention describes a novel method for the production of  
CC ketocarotenoids which involves culturing genetically modified organisms  
CC having modified ketolase (KLA) activity. The starting microorganisms  
CC produce carotenoids (naturally or by genetic supplementation), and are  
CC specifically microorganisms (especially bacteria, yeasts, algae or fungi)  
CC or plants. Microorganisms and plants involved in the method could include  
CC Escherichia, Flavobacterium, Nostoc, Synochocystis, Hansenula, Fusarium,  
CC Dunaliella, Rhodospirillum rubrum, Rhodospirillum rubrum, Rhodospirillum rubrum,  
CC Solanaceae, Lamiaceae, Acacia, Calendula, Gentiana, Helianthus, Linum,  
CC Rhododendron, Spartium and zinnia. The modified microorganisms  
CC additionally show elevated hydroxylase and/or beta-cyclase activity. The  
CC ketocarotenoids described in the invention include astaxanthin,  
CC adonirubin, canthaxanthin, adoxanthin and zeaxanthin. Ketocarotenoids are  
CC natural antioxidants and pigments, especially useful as pigmenting  
CC additives in animal feed, specifically feed for trout, salmon or shrimps.  
CC Ketocarotenoid-producing genetically modified organisms can be used as  
CC feedstuffs or foodstuffs, in the production of ketocarotenoid-containing  
CC extracts or for producing feed or food supplements. The process provides  
CC large amounts of ketocarotenoids having a low content of hydroxylated by-  
CC products, especially in the case of astaxanthin, preferably due to  
CC expression of at least one nucleic acid encoding hydroxylase and/or beta-  
CC cyclase. DNA encoding the whole primary ketolase sequence from Nostoc sp.  
CC strain PCC7120 was isolated, amplified by PCR and used to produce a

CC plasmid pNOSTF-G. A plasmid pMCL-Crt-YIBZ/idi/gps, for the synthesis of  
CC zeaxanthin in Escherichia coli, was constructed in 3 stages via the  
CC intermediate stages pMCL-CrtYIBZ and pMCL-CrtYIBZ/idi, using the high  
CC copy number plasmid vector pMCL200. Escherichia coli strain TOP10 was  
CC transformed with the plasmids pNOSTF-G and pMCL-Crt-YIBZ/idi/gps to give  
CC carotenoid producing strain.  
XX  
SQ Sequence 4624 BP; 1026 A; 1243 C; 1270 G; 1085 T; 0 U; 0 Other;  
Query Match 27.2%; Score 1900.6; DB 12; Length 4624;  
Best Local Similarity 66.5%; Pred. No. 0;  
Matches 2787; Conservative 0; Mismatches 1389; Indels 13; Gaps 4;  
QY 2405 GCTACGGATGTCAGCGATTACAGGCGCAGCTGAGCGCGCGAGCGGTGGCAGCGTGCCTCC 2464  
DB 12 GCAGCGCATGGCGAAATCCAGACAGCCCTTCGTTTGGCAGGGGCGACCATGGCGCGTGC 71  
QY 2465 CGACATCTCGAGCAGCGCGTGTGCGACGACGAGCTGCTGGCGGAGGCGACCTGATG 2524  
DB 72 CGATATCATTTAGCAGCGATTATGTG-CACCGGTGACCTGTCTTAAGTGGAGCGGCTATG 130  
QY 2525 CGCAGCAATACGATGTGATTTTGGTGGTGTGCTGGAATGGGAAATGGCTTGTATCGCTG 2584  
DB 131 CAACCGCATTTATGATCTGATTTCTGCGGGGCTGAGCTCGCAATGGCTTATCGCCCTG 190  
QY 2585 GGTCTGCGTCAATTGCGACGCAACTGAAATGCCCTGTTGCTGGAGAGGATCGGATCCG 2644  
DB 191 GGTCTTCAGCAGCAGCAACCTGATATGCTGATTTTGTCTTATCGACCGCGCACCCAGGCG 250  
QY 2645 GCAGCAATCATACCTGCTGCTTTCATCAGCGGATCTCAGCGCGCAACACTTCGCTGG 2704  
DB 251 GCGCGGAATCATACGCTGCTCATTTCCACCAATGATTTGACTGAGAGCAACATCGTTGG 310  
QY 2705 CTGCAACCGCTGATTTACCGTGGTGGTTCAGGTGATCAGGTGCGTTTTCTCGCTGCGC 2764  
DB 311 ATAGCTCGCTGGTGGTTCATCAGTGGCGGCTGATCAGTACGCTTCCACACAGCGCT 370  
QY 2765 CGCAATCTGGACGGGGATTTATGTTCCATCGCATCGAGCGGATTTTGGCCGCCATCTTTAC 2824  
DB 371 GGTAACTGAAACAGCGGCTACTTTTGTATTTACTTCTCAGCGGTTTCGCTGAGGTTTTACAG 430  
QY 2825 GCGCGATGGGTGACGATCTGTGGACAACACAGCGGTACACAGGTAAACCCAGCGAG 2884  
DB 431 CGACAGTTTGGCGCGCACTTGTGGATGATACCGCGGTGCGAGAGTTTAATCGGAATCT 490  
QY 2885 GTGACGCTGGCGGATGGCGGTGAACCTTGTGCGCAAGTGGTGAATGATGCTGCGCGCTG 2944  
DB 491 GTTCGGTTGAAAAGGGTCAAGTTATCGTGTCCCGCGGTGATTTGACGGCGGGGTTAT 550  
QY 2945 CAGCGACGCCACATCTCGAGCTGGGTTATCAGGTGTTTCTTGGACAGAGTGGCAGCTG 3004  
DB 551 GCGGCAAAATTCAGCACTGAGCGTGGGCTTCCAGGCGTTTATTTGGCCAGGAATGGCGATTG 610  
QY 3005 GCGCAGCGCAGCGCTCGCAGCGCGGATCTGTGATGATGCCACCGTGCATCAGCAAGCG 3064  
DB 611 AGCCACCCGCGATGGTTATGCTCTCCCATATCATGGATGCCACCGTGCATCAGCAAAAT 670  
QY 3065 GGTATCTGTTTGTCTACACGCTGCGCTGCGCTGAGCGCGGATCGGCTATTGATGAAGATAC 3124  
DB 671 GGTATCTGTTTGTCTACACGCTGCGCTGCGCTGCGCTGCGCGACGAGATTGTTAATGAAGACAG 730  
QY 3125 CATTAAGTTAACAGCGCGCTGGCGGAGAACACCGTCTGTCAGACATTCGCCGACTAT 3184  
DB 731 CACTATATTGATAATTCGACATTTAGATCCTGAATCGCGCGGCAAAATATTTGCGACTAT 790  
QY 3185 GCCAATCAGCAGCGCTGGACGCTGAGTACGCTGCTGCTGAGAGCAGCGCATATTACCG 3244  
DB 791 GCGCGCAACAGGGTGTGAGCTTCAGACACTGCTGCGAGAGAACAGCGCGCTTACCC 850  
QY 3245 ATTACCTCTGAGCGCAACATCGATCGATTTCTGGCAACAGCAGCGCGGCGCAAGCGTGCAGC 3304  
DB 851 ATTACTCTGTGCGGCAATTCGCGACCGCATTTCTGG---CAGCAGCGCGCTTCTGCTGCTGCT 907

QY 3305 GGCCTCGCGCGCGGCTGTTTCATGCCACACCGGTTACTCCTTGCCGTCGCGCGTGGCG 3364  
Db 908 GGATTAGGTGGCGGCTGTTTCATCTACCAACCGGCTATTCTACGCGCTGGCGGTTGCC 967  
QY 3365 CTAGCGGAGTTGGTAGCAGCGCTGTTGCCACCGATGCCCTCACGCTCAGCCAACTATC 3424  
Db 968 GTGGCGGACCGGCTGAGTGCACTTGATGTCTTTAGCTCGGCGCTAAITCACCATGCCAAT 1027  
QY 3425 GAAACGCTTTGCCGCTCAGCAGTGGCGGCAACAGCGATTTCCTGCTCTGCTAAACCGCATG 3484  
Db 1028 AGCAATTTGCCGAGCGCTGGCAGACGAGGCTTTTCCGATGCTGATCGCATG 1087  
QY 3485 CTGTTTTGGCGGTAAGCCGAGCAGCGCTGGCGGTGANGAAGCTTTTACCGGCTC 3544  
Db 1088 CTGTTTTAGCCGGACCGCGCGATTCAACGCTGGCGGTTTATGACGCTTTTATGTTT 1147  
QY 3545 GATCGCGGTTAATTAGCGCTTTTACGCGGGGAACTTGGCGCTCGCGCTGTTGAATCT 3604  
Db 1148 CCTGAAGATTTAATTGCGCGTTTATGCGGGAAACTCACGCTGACCGATCGGCTACGT 1207  
QY 3605 ATTCTGCGGCAAGCGCGCTGCCATCGGTGAAGCGCTCGCGCGCTGTTGAATCT 3664  
Db 1208 ATTCTGAGCGGACGCGCTGTTCCGTTATTAGCAGCATTCGAAGCCATATGAGACT 1267  
QY 3665 GTCGAACCAAGGAAGAAATGAAACGCACTTATGTATTGGCGCAGGCTTTGGCGCC 3724  
Db 1268 CATCGTTAAAGAGCACTACATGAACCAACTACGCTAATTTGTCGACGCTTCGCTGCC 1327  
QY 3725 TGGCGTGGCGATTGCGCTGGAAGCGGCGGATACCAACCACTTACTCGAGCAGCGG 3784  
Db 1328 TGGCACTGGCAATTCGCTACAACTGCGGGATCCCGCTCTTACTGCTTGAACACGCTG 1387  
QY 3785 ACAACCGCGGCGAGCGCTATGTGTTGAGGACAGTGGCTTTACTTTCGATGCGGAC 3844  
Db 1388 ATAAACCGCGCGTGGGCTTATGCTACGAGATCAGGGTTTACTTTGATGACGCGC 1447  
QY 3845 CCACGCTGATCCGATCCAGCGCATCGAAGAGTTGTTACGCTGGCAGGAAATCGC 3904  
Db 1448 CGACGGTTATCCGATCCAGTCCAGTGAAGAACTGTTGCACTGGCAGGAAACAGT 1507  
QY 3905 TCAGCGATTACTCGAGCTGATCGGTTAAGCGCTTCTATCGCTGCTGGAGATG 3964  
Db 1508 TAAAGAGATGTCGAACCTGCTCGCGTTACCGCTTTTACCGCTGTTGGAGTCA 1567  
QY 3965 GCAACAGCTGATTACGACAATAATCAGCGCTGCTGGAGCAGACATCGCACGTTCA 4024  
Db 1568 GGAAGTCTTTAATTACGATACGATCAACCGCGCTCGAAGCGCATTCAGCAGTTA 1627  
QY 4025 ATCCGCAAGATGTAAGGCTATCGTCAATTTCTTGCTTATTCAGTGAAGTATTAGAG 4084  
Db 1628 ATCCCGCGATGTCGAAGGTTATCGTCAGTTTCTGGACTATTTCAGCGCGGTTTAAAG 1687  
QY 4085 AGGTTATCTGAACCTCGGCAAGTGGCGCTTCTGCGAGTGGTGACATGCTGGCGTC 4144  
Db 1688 AAGGCTATCTAAGCTCGGTACTGTCCTTTTATCGTTTCAGACATGCTTCGCGCG 1747  
QY 4145 CGCGCAGTTGGCAGCTCTGCAAGATGCGCAGCTCTACAGCATGTTGGCGCAATTTA 4204  
Db 1748 CACTCACTGGGAACTGCGAGCATGGAAGCGTTTACAGTAAGTTGCGCAGTTACA 1807  
QY 4205 TTCAGACGATCATCTGCTCAGGCGTTTCTTCCACTCATTTGCTGGCGGTAATC 4264  
Db 1808 TCGAAGATGAACATCTGCGCCAGGCGTTTCTTCCACTCGCTGTTGGTGGCGGCAATC 1867  
QY 4265 CTTTTCAGCGTCTATACCTTAATTCATGCGCTGAGCGTGAATCGGCGTGT 4324  
Db 1868 CCTTCGCCACCTCATCTTTATACGTTGATACACGCTGAGCGTGAATGGGCGTCT 1927  
QY 4325 GGTTCGCGCGCGGCAACCGCGCTGTCAGGCGCATGGCGACTGTTTCGAGGACT 4384  
Db 1928 GGTTCGCGCTGGCGGCAACCGCGCATTAAGTTCAGGGATGATAAGCTGTTTCAGGATC 1987  
QY 4385 TGGCGCGGAGCTGTTACTGAATGCGGAAGTGAAGCGCTGGAAACCGACGCGCAATCGCA 4444

Db 1988 TGGTGGCGAAGTCTGTGTTAAACCGCAGAGTCAGCCATATGGAACACGACGGAACAGA 2047  
QY 4445 TTAGCGCGTTCAGTTAGAGGCGGACAGCTTCGATGCGCGCTGTGCTCTCCAAATG 4504  
Db 2048 TTGAAGCGGTGCATTTAGAGGACGCTCGAGGTTCTGACGCAAGCGCTGCGGTCAAATG 2107  
QY 4505 CCGAGTGGTGCATACCTTACGACAAATGCTTCGCCACCACTCCGCTGGCAATGAACGTG 4564  
Db 2108 CAGATGTGTTTACATACCTTATCGGACCTGTTTAAAGCAGCACCTTGCCTGTAAGCAGT 2167  
QY 4565 CGACATCGCTGAAGGTTAAGCGGATGAGCAACTCGCTGTGTTGTAATCTTATTTTGGCCTGA 4624  
Db 2168 CCAACAAACTCGACACTAAGCGCATGATTAATCTCTGTTTGTCTCTATTTTGGTTTGA 2227  
QY 4625 ATCAGCGCATGAACAGCTCGCGCACCAACCGCTGTTTGGCCGCGTTATCGTGAAT 4684  
Db 2228 ATCACCATCATGATCAGCTCGCGCATCACAGGTTTTCGCGCCGCGTTTACCGCGAGC 2287  
QY 4685 TGATCGATGAGATTTTCAACAGCAGCAGCTGGCAGACGATTTTTCATCTTTACCTGCAAG 4744  
Db 2288 TGATTGACGAATTTTAAATCATGATGGCTTCGACAGGACTTCTCACTTTATCTGCACG 2347  
QY 4745 CGCCTCGCAGCAGCGATCCGCTCGCTGCGACCGCGCGCTGCGCAGCTTTTATGTTAG 4804  
Db 2348 CGCCTGCTGTCAAGGATTCGCTACTGCGCGCTGAAGGTTGCGGAGTTTACTATGTTGG 2407  
QY 4805 CGCGGTGCGGCATCTCGGCACCGCTGACATCGACTGGCAACAGAGGACCGCGCTTGC 4864  
Db 2408 CGCGGTGCGGCATTTAGGACCGCGAACTCGACTGACGTTGAGGGGCAAACTAC 2467  
QY 4865 GCGATCGAATTTTGTCTTATCTGGAGCAGCACTACATGCGCGGATTAAGTCAGCAATPAG 4924  
Db 2468 GCGCGGTATTTTGGCTTACCTTGAGCAGCATTAATGCTGCTTACGAGTCACTGCTGG 2527  
QY 4925 TGACACACAGAAATGTTTACGCGCTTGTATTTTCGCAACAGCTGCTGATGCTTCACTGCT 4984  
Db 2528 TCACGACCGGATGTTTACGCGCTTGTATTTTCGCAACAGCTTAAATGCTTATGCTGCT 2587  
QY 4985 CGCGGTTCGCTGGAGCGGATTTTACGCAAGCGCTGCTGCTCGCGCGCATTAACCGCG 5044  
Db 2588 CAGCTTTTCTGTTGAGCGCGCTTCTTACCGAGCGCTGCTTTCGCGCGCATTAACCGCG 2647  
QY 5045 ATCGCATATCAGCAATCTCTATCTGTTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5104  
Db 2648 ATAAACCATTAATCTCTACTCTGCTGCGCGCAGGCAACGCTGCTGCTGCTGCTGCTGCT 2707  
QY 5105 CGCGGTGATCGGTTTCGCGCAAGCGCCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5164  
Db 2708 CTGCGCTCATCGCTCGGCAAGCGCAGCAGGTTTGTGCTGAGGATC---TGATTT 2764  
QY 5165 GAATCGACAGCTTTTACTTGTAGCAAGTAAACGCAACCATGCGGCTGCTGCTGCTGCTGCTGCT 5224  
Db 2765 GAATATCCGCTGCTTACTCAATCATCGCTGCAAAAGTGGAGTTGCTGCTGCTGCTGCTGCT 2824  
QY 5225 CGCCACCGCGCGCAAGCTGTTTGTATGACCGCGCGCGCAGCAGCTGCTGCTGCTGCTGCTG 5284  
Db 2825 TCGACAGCTCAAGTATTTTGTATGCAAAACCGCGCGCAGCAGCTGCTGCTGCTGCTGCTGCT 2884  
QY 5285 GTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5344  
Db 2885 CTGTTGCGCGCATTTGTGACGATGTTTGTGACGATCAGACGCTGCTGCTGCTGCTGCTGCTG 2944  
QY 5345 GCATCGCTCGAAGCAGCGCAGGACGCTGATGCAAGCTTCTGCAATCTGCAAAATGGAACCGCGCGCG 5404  
Db 2945 GCTGCTTACAAACCGCGCGCAACGCTGCTGATGCACTTGTGATGAAACCGCGCGCGCG 3004  
QY 5405 CTACAGCGCGCGCAGCATGATGAACCGCGCTTTAGGGCGTTTACGAGGATGCTGCTGCTGCTG 5464  
Db 3005 CTATCAGGATCGCAGATGCAACCGCGCTTTTTCGCGCTTTTTCAGGAGTGGCTGCTGCTG 3064  
QY 5465 TCACAGCTGCGGCAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5524

Db 3065 TCATGATATCGCCCGCTTACGCGTTGATCATCTGGAAGGCTTCGCCATGATGTACG 3124

Qy 5525 CAAACGAATACGCGAGCTTCGATGACACGCTGCGTTACTCTATACGTCGCGGCGT 5584

Db 3125 CGAAGCGCAATACAGCAACTGATGATACGCTGCGCTATTGCTATACGTTGCAAGCGT 3184

Qy 5585 GGTGCGTTGATGATGCGGCGTAAATGCGCGTTCGCGGACGAGCGGTCTGATCAGCG 5644

Db 3185 TGTGCGCTTGAATGATGCGCAATCATGCGCGTTCGCGGATACGCCACGCTGGAACGCGC 3244

Qy 5645 CTGCGATTTAGACTGCGCTTCAGCTCACTAACATTGCGCGACATTTGATGAGATGC 5704

Db 3245 CTGTGACCTTGGCTGGCATTTGATGACCAATATTGCTCGGATATTGTGCGAGTGC 3304

Qy 5705 CGAAATGTCGCTGCTATCTCGCGCAATCTGCGTTCGATCAGCGGGAATTACGCGCGA 5764

Db 3305 GCATCGGCGCGCTGTTATCTCGCGCAAGCTGGCTGGAGCATGAAGGTCTGAACAAAGA 3364

Qy 5765 TAGCTGATGACCGCAACATGCTGACGCGTTCGCTCACTGGCAGCGCGTTTATGTGC 5824

Db 3365 GAATTATGCGGCACCTGAAACCGTCAGCGCTGAGCGCTATCGCCCGCTGTTGGTGA 3424

Qy 5825 GGAGCGGAACCTATTATCACTCGCGCGATCCGGTTTACCGGTTTACCGCTCGGCTC 5884

Db 3425 GGAAGCAGAACCTTACTATTGCTGCCACAGCGGCTGGCAGGGTTGCCCTCGGTTTC 3484

Qy 5885 GCGGTGGCGATCGCTACGCTCGCGGCTTTTATCGGAAATTTGGCGTCAAAAGTTCAACA 5944

Db 3485 CGCTGGGCAATCGCTACGCGCAAGCAGGTTTACCGGAAATAGGTGTCAAAAGTTGAACA 3544

Qy 5945 CGCCGCTGACAGCTGGATTCAGCGCAGCGCACCAGTAAAGGTGAATACTGGCGT 6004

Db 3545 GCGCGGTGACGAAGCTGGATGACGCGCAGTCAACAGCACCGCCGCAAAATTAACGCT 3604

Qy 6005 GCTGTGAAAGCGGAGGTTTGGCGATCACTTCGCGTGTCTCGTCTGTAACCGCTCC 6064

Db 3605 GTGCTGGCGGCTCTGCTCAGCGCTTACTTCCGGATGCGGCTCATCTCCCGGCC 3664

Qy 6065 GCGTGTCTGTGCGAGCTCTCTGTTGATTTTACGTCGCTGACGCTGGCGCAGCGTGT 6124

Db 3665 TCGCATCTCTGCGACGCGCGC-----TCTAGCGCCATGCTTTTCCGGAGCTGCGC 3718

Qy 6125 TCGAGTTATTACGCGTGGCGGTAGAGAAACCAACGACACGCGAGCTTCAGCGCG 6184

Db 3719 TGAAGTTTTCAGCGGCGCGCATAGAGAAAGCAACCAACCACTTCTTTTGCCC 3778

Qy 6185 CGCACGCGATGATGATGCGGTGCGCATATAGCGCTTAAGATAGCTTTTGGCGGG 6244

Db 3779 CTGACGCGTGTATGATACGCTGGCGCATATACACCGTTTGGAGTAGCCCTTGTGTGA 3838

Qy 6245 ATATAGCGGAACGCGCGTTGATGACACGCGCATCTGACCATGAAGTAGAGCGG 6304

Db 3839 ATATAGCGGAATGGCGCAAGTTGATGACACGCGCTGTCACCATAAATAGAGTAAT 3898

Qy 6305 CGTAGCTGTCTATTCGCGCAACATCCACTGCGGCGCATGCTTGCACACGCA 6364

Db 3899 CCATAGCGGTATACCTCGGCAATCCACTGAGCGGCCACATCTCTGTACTGCGCCAGA 3958

Qy 6365 TAAATCAGCAACATCCAGTACCGCAACACACCGCAATAAGATCGTTGAGCTCAAC 6424

Db 3959 TAAATCAGCAGNTGATGATGAGCAAAACCAACGCGCATAAAGATCGTTAATCTCAAC 4018

Qy 6425 TTACCGCTGTGCGGTTCATGTTGCGACAGATCCAGCCCGCATCCCAACCGTGCATGATG 6484

Db 4019 GCACCTTTACGCGGTTTATGATGTGAAGATGCCATCCCAACCCCGCGGTGCATGATG 4078

Qy 6485 TATTATGCGACAGCGCGCTAGATTTCATCACCACACCGTTGCCAGCAAGATAGC 6544

Db 4079 TATTGTGTGCGAGTGCAGCAATCACTTCCATGCCAATCACCGTAAACGAAACGATCAG 4138

Qy 6545 ACCTTCCATAACACAGAGCATTTGCTGCTCATTTGTGGAAGGGAAGTA 6593

Db 4139 GCATTCAAATCCACAAATATTTCTCCGTTAGACGCTGTGGACGA 4187

RESULT 8  
ABZ69177  
ID ABZ69177 standard; DNA; 7494 BP.  
XX AC ABZ69177;  
XX DT 11-AUG-2003 (first entry)  
XX Vector construct p8Scarola4 coding sequence.  
XX Biosynthesis; enzyme; antibacterial; protozoacide; herbicide; isoprenoid;  
KW inhibitor; immunostimulant; gene; ds.  
XX OS Unidentified.  
XX OS Synthetic.  
XX FN DE10201458-A1.  
XX PD 17-OCT-2002.  
XX PF 16-JAN-2002; 2002DE-01001458.  
XX PR 11-APR-2001; 2001DE-01018166.  
PR 22-JUN-2001; 2001DE-01030236.  
PR 09-NOV-2001; 2001DE-01055084.  
XX (BACH/) BACHER A.  
PA (ROHD/) ROHDICH F.  
XX Adam P, Amsingler S, Bacher A, Eisenreich W, Hecht S, Rohdich F;  
WPI; 2003-185556/19.  
XX New proteins involved in isoprenoid biosynthesis, useful in screening for  
inhibitors, also new intermediates, potential therapeutic agents, nucleic  
acids and antibodies.  
XX Disclosure; Page 45-47; 78pp; German.  
XX The present invention relates to proteins capable of the enzymatic  
conversion of 2C-methyl-D-erythritol-2,4-cyclodiphosphate to 1-hydroxy-2-  
methyl-2-butenyl-4-diphosphate. The proteins are useful in screening a  
chemical library for inhibitors of isoprenoid biosynthesis. Cells and  
organisms that express them are useful for in vivo/in vitro production of  
isotopically labeled biosynthetic intermediates and products of the  
mevalonate-independent pathway of isoprenoid synthesis. Such  
intermediates and products are used to screen for genes, enzymes or  
inhibitors of isoprenoid/terpenoid biosynthesis, potentially useful as  
antibiotics, antimetabolites and herbicides and as immunostimulators,  
particularly for treating bacterial infection. Antibodies against the  
intermediates are useful for detecting pathogens, in body fluids. The  
present sequence is a coding sequence used in the exemplification of the  
invention  
XX Sequence 7494 BP; 1722 A; 1935 C; 2026 G; 1811 T; 0 U; 0 Other;  
SQ  
Query Match 23.1%; Score 1618.8; DB 8; Length 7494;  
Best Local Similarity 66.0%; Pred No. 0;  
Matches 2375; Conservative 0; Mismatches 1217; Indels 6; Gaps 2;  
Qy 2490 CAGCAGCAAGTGTGTCGCGGAGCGGACCTGATCGCAGCAACATACGATGTGATTTGG 2549  
Db 3213 CACCGGTGAGCTGTCTTAAAGTGGAGCGGCTATGCAACCGCATTTATGATCTGATCTCG 3272  
Qy 2550 TCGGTGCTGGAGTGGCGAATGGCTTGATTCGCTCGCTCGCTCAATTGAGCCACAC 2609  
Db 3273 TGGGGGCTGGAGTTCGGAATGGCTTATGCGCCCTCGCTTTCAGCAGCAGCAACCTGATA 3332  
Qy 2610 TGAATGCTGTGTGCTGGAGAGCGATGCGCATCCGCGAGGCAATCATACCTGCTGTTTC 2669  
Db 3333 TCGGTATTTTGTCTTATCGACCGCCACCCAGCGGGGGGATCATACGTGTCATTTTC 3392

QY 2670 ATCAGCGATCTCAGCGCCGAAACAACCTTCGTGCTCGCAACCGCTGATTACCGTGCTT 2729  
Db 3393 ACCACGATGATGACTGAGACCAACATCGTTGGATAGCTCCGCTGGTGTTCATCACT 3452  
QY 2730 GGTCAAGTTATCAGGTGGGTTTCTCGCTGCGCGCAATCTGACCGGGGATATTGTT 2789  
Db 3453 GGGCCGACTATCAGGTAAGCTTTCACACGCCGTCGTAAAGCTGAACAGCGGTACTTTT 3512  
QY 2790 CCATCGCATCAGCGCATTTTGGCCGCCATCTTTTACGCGCGAGTAGGGTGACGATCTGTGA 2849  
Db 3513 GTATTACTCTCAGCGTTTCGCTGAGGTTTACAGCGACAGTTTGGCCGCACTTGTGA 3572  
QY 2850 CAAACACAGCCGTAACAAGGTAAACCAACGACGTCGCTGCGCGATGSCGCTGAAC 2909  
Db 3573 TGGATACCGCGTCGACAGGTTAATGCGGAATCTGTTCCGTTGAAAGGTCAGTTA 3632  
QY 2910 TTGCTGCGCAAGTGTGATGTTGATGTTGCGGCTCGACGCCGACACATCTGCACTGG 2969  
Db 3633 TCGGTGCCGCGGGTGTGACGGCGGGGTTATGCGCAATTCAGCACTGAGCGTGG 3692  
QY 2970 GTTATCAGGTGTTCTTGACAAAGTAGTGGCAGCTGGCGAGCGCCGACCGCTGACGAGC 3029  
Db 3693 GCTTCCAGCGGTTTATGTCAGGAATGCGGATGAGCCACCGCATGGTTTATCGTCTC 3752  
QY 3030 CGATCTGATGATGCCACCGTCGATCAGCAAGCGGTTATCGTTTGTCTACACGCTGC 3089  
Db 3753 CAAATATCATGATGCCACGGTCGATCAGCAAAATGGTTATCGCTTGTGTACACCTGC 3812  
QY 3090 CGCTCAGCGCCGATCGGCTATTGATGAAAGATACCAATPACGTTAACCGACCGCGCTGG 3149  
Db 3813 CGCTCTCGCGCAGCAGATGTTAATGAAAGACAGCACATATATGATAATGCGACATTAG 3872  
QY 3150 CCGAGAACACCGCTCGTCAGCACATCGCGATCTGCCAATCAGCAAGCTGAGACGCTGA 3209  
Db 3873 ATCTGAAATCGCGCGGCAAAATATTGCGACTATGCGCGCAACAGGGTTGGCAGCTTC 3932  
QY 3210 GTACGCTGCTGCGTGAAGACGCGCATATTACCGATTAACCTGAGCGGCAACATCGATC 3269  
Db 3933 AGACACTGCTGCGAAGAACAGGGCGCTTACCCATTAATCTGTCGGSCAATGCCGAGC 3992  
QY 3270 GATCTGGCAACAGCAGCGCGGCAAGCGTGACGCGGCTGCGCGCGGGCTGTTTCATG 3329  
Db 3993 CATTTCTGG---CAGCAGCGCGCCCTCGCTGTAGTGAATTAAGTCCGCTCTGTTCCATC 4049  
QY 3330 CCACACCGGTTACTCTTTCGCTGCGCGTGGCGCTAGCGAGTTGGTAGCAGCGCTGT 3389  
Db 4050 CTACCAACCGGCTATTCTGCGCTGCGGTTGCGGTGCGGTGCGGACCGCTGAGTGCACTTG 4109  
QY 3390 TGCCCAACCGATGCCCTCAGCTCAGCCAAACATATCGAACCGTTTGGCCGTCAGAGTGGC 3449  
Db 4110 ATGCTTTTACGTCGGCTCAATTACCAATGCCAATTACGCAATTTTGGCCGAGCGCTGGC 4169  
QY 3450 GCGAACAGCGATTTTTCGCTCTCTAAACCGCATGCTGTTTGGCGCGTAAGCCGACG 3509  
Db 4170 AGCAGCAGGGCTTTTTCGCGATCGAATCGCATGCTGTTTGTAGCCGACCGCGGAT 4229  
QY 3510 AGCGTGGCGGTGATGCAACGTTTATACCGGCTCGATGCGGGTTAAATTAGCCGCTTTT 3569  
Db 4230 CAGCTGGCGGTTATGACGGTTTATATGTTTACCTGAAGATTTAATTTGCCCGTTTTT 4289  
QY 3570 ACSCCGGGCAACTGCGCTGCGGATAAACCGCGGATTTGTGGCGAAGCCGCGCGTGC 3629  
Db 4290 ATGCGGAAACATCAAGCTGACCGCTAAGTATTTCTGAGCGGCAAGCCGCTGTTC 4349  
QY 3630 CCAATCGGTGAAGCGCTGCGCGCTGTTGAATTTCTGTGCAACCAAGGGAAGAAATGAA 3689  
Db 4350 CGGTAATTAGCAGCATTTCAAGACCATTAATGACGATCATCTGTTAAAGCGCATACATGAA 4409  
QY 3690 ACSCACTTATGTGATGGCGAGGCTTTGCGGCTGCGGCTGCGGATTCGCTGCAAGC 3749  
Db 4410 ACCAACTACGTTAATGGTCAGGCTTCGGTGGCTGGCACTGGCAATTCGTCTACAAGC 4469

QY 3750 GGGCGGCATACCAACACCTTACTCGAGCAGCGCGCAAAACCGGCGGACGCGCCTATGT 3809  
Db 4470 TCGGGGATCCCGCTCTTACTGCTTGAACAACCGTGAATAACCGCGGTCCGGCTTATGT 4529  
QY 3810 GTTTAGGAGCAGTGGCTTTTACCTTCATGCGCGGACCAACCGTGTATCACCGATCCGAGCGC 3869  
Db 4530 CTACGAGGATCAGGGGTTTACCTTTGATGACGCGCCGAGCGGTTATCACCGATCCGAGTGC 4589  
QY 3870 CATCGAAGAGTGTTCACGCTGCGAGGAAATCGCTCAGCGATTAAGTCTGAGCTGATGCC 3929  
Db 4590 CATTTGAAGAACTGTTTGCATCGCAGGAAACAGTTAAAGAGTATGTGAACTGTCTGCC 4649  
QY 3930 GGTAAACCCCTTCTATCGCCTGCTGCGGAAGATGCAAAACAGCTGATTAACGACAATAA 3989  
Db 4650 GGTACCGCTTTTACCGCTGTGTTGGAGTCAAGGAGGCTCTTTAAATACGATAACGA 4709  
QY 3990 TCAGCCGCTGCTGAGCAGCAGATCGCACTTCAATCCGCAAGATGTAGAAGCTATCG 4049  
Db 4710 TCAAAACCGCTCGAAGCGCAGATTCAGCAGTTTAAATCCCGCATGTGCAAGGTTATCG 4769  
QY 4050 TCAATTTCTTGCCTATTCAAGTGAAGTATTAGAGAGGTTTATCTGAAACTCGGCACGCT 4109  
Db 4770 TCAGTTTCTGACTTATTCAGCGCGGTGTTTAAAGAGGCTATCTAAAGCTCGTACTGT 4829  
QY 4110 GCCGTTTCTGAGGTGCTGACATGCTGCGCTGCGCCGCGAGTTGGGAAGTCTGCAAGC 4169  
Db 4830 CCCTTTTTATCGTTTCAGAGACATGCTTCGCGCGCACTTCAACTGGCGAACTGCAAGC 4889  
QY 4170 ATGCGCAGCGCTPACAGCATGGTGGCAAAATTTATTCAGACCATCATCTGCGTCAAGC 4229  
Db 4890 ATGAGAAAGCGTTTACAGTAAGTTGCGCATTCATCGAAGATGAACATCTGCGCCAGC 4949  
QY 4230 GTTTTCTTCCACTCATTTGCTGTGGCGGTAAATCTTTTGGCAACGCTCATCGATCTATAC 4289  
Db 4950 GTTTTCTTCCACTCGCTGTTGGTGGCGCAATCTCCTTCGCCACCTCATCTCAATTTATAC 5009  
QY 4290 CTTAATTCATGCGCTGAGCGTGAAATGGGGGCTGTGTTTCCGCGCGCGGCAACCGCGC 4349  
Db 5010 GTTGATACACGCTGAGCGGTGAGTGGGGGCTGTGTTTCCGCTGCGCGCACCGCGC 5069  
QY 4350 GCTGTGTCAGGGCATGGCGCATGTTTCAGAGACTTGGGGCGGAGCTGTTTACTGAATGC 4409  
Db 5070 ATTAGTTCAGGGATGATAAAGCTGTTTCAGACTCTGGGTGCGCAAGTCTGTTTAAACGC 5129  
QY 4410 CGAAGTAGCAGCTGGAACACGCGCAATTCGCAATTAAGCGGCTTACGTTAGAGGCGG 4469  
Db 5130 CAGAGTCAGCCATATGGAACACGACAGGAAACAAGATTGAAGCCGTGCAATTTAGAGCAG 5189  
QY 4470 ACAGCGTTTCGATGCGCGCTGTGGCTCAATGCGGAGCTGTGATCTACTACGACAA 4529  
Db 5190 TCGCAGGTTCTGACGCAAGCCGTCGCTCAATGCGAGTGTGTTTATCTATCTGCGGA 5249  
QY 4530 ACTGCTTCGCCACCATCCGCTGGCAATGAACCGTCCGACATCGTGAAGCGTAAGCGCAT 4589  
Db 5250 CCGTTTAAGCCAGCACCCCTGCGCGGTTAAGCAGTCCAAACAACTGACAGCTAAGCGCAT 5309  
QY 4590 GAGCAACTCGCTGTTTGTACTTATTTTGGCTGATCAGCGGATGCAACAGCTGCGGCA 4649  
Db 5310 GAGTAACCTCTCTGTTGTGCTTATTTTGGTTTGAATCACCATCAGATCAGCTGCGGCA 5369  
QY 4650 CCACACCGCTGTTTGGCGCGGTTATCTGTGAGTTGATCGATGAGATTTTCAACAGCAG 4709  
Db 5370 TCACACGGTTGTTTTCGCGCGGTTACCGCGAGCTGATTGACGAAATTTTAAATCATGA 5429  
QY 4710 CCAGCTGGCAGACGATTTTTCATTTAAGTCTGAGCTGCGGCTGCGAGCAGTCCGCT 4769  
Db 5430 TGGCCCTCGCAGAGGACTTCTCACTTATCTGACGCGGCTGTGTGTCAGGATTCGCTACT 5489  
QY 4770 GGCACGCGCGCTGCGCAGCTTTTATGTTTGTAGCGCGGTGCGCATCTCGGCAACCGC 4829  
Db 5490 GCGCGCTGAGAGTTGCGGAGTTACTATGTTTGGCGCGGTGCGCATTTAGGCAACCGC 5549  
QY 4830 TGACATCGACTGGCAACAGGAGGACCGCGCTTGGCGGATCGAATTTTGTGTTTCTGGA 4889

Db	5550	GAACTCGACTGGAAGTGGGGGCAAAACATACGGACCGTATTTTTCGCTACTTGA	5609	Db	6267	CAGGTTTACCGAAATAGGTGTCAAAGTTGAACAGCGCGTCAAGACCTGGGATCAG	6686
Qy	4890	GCAGCATACATGCGCGGATTAAGTACAGCAATAGTAGACACACAGAAATGTTTACGCGGTT	4949	Qy	5970	CGGACGACACACAGTAAAGGTGAAAACTGGCGTGTGGTGAAGGGCAGGTTGGCG	6029
Db	5610	GCAGCATACATGCGCTTACGGATCAGCTGGTTCAGCACCAGATGTTTACGCGGTT	5669	Db	6687	CGGAGTCAACAGCACGACCGCCGAAAAATTAACGCTGCTGTGGCGCCTCTGGTCAGGCC	6746
Qy	4950	TGATTTTCGGACACGCTGCATGCCATCACGGCTCGGGGTTTTCGCTGGAGCGATTTT	5009	Qy	6030	ATCACTTCGGTGTGTCGTCTCTGAAACGGGTGCGGCTGTCTGGCAGCGTCTTC	6087
Db	5670	TGATTTTCGGACCAAGCTTAATGCTTATCATGCTTTCAGGCTTTTCTGTGGAGCCGCTTCT	5729	Db	6747	CTTACTTCCCGATGCGGCTCATCTCCCGCCCTGCGCATCTCTGGCAGCGCCGC	6804
Qy	5010	GACGCAAGCGCTGTGTTCCGCGCATAAACCGCATGCCGATATCAGCAATCTCTATCT	5069	RESULT 9			
Db	5730	TACCCAGAGCGCTGTGTTGCGCGCATAAACCGGATTAACCAATTAATCTTACT	5789	ID	ABZ69178	standard; DNA; 8547 BP.	
Qy	5070	GCTGGGTGCGGTAGCATCCAGGCGCGGCGGTGCCCGCGTGATCGGTTCCGCCCAAGGC	5129	XX	ABZ69178;		
Db	5790	GCTGGCGCAGCAGCATCCCGCGCAGGCATCTCTGGCGTCATCGGCTCGGCAAGC	5849	AC	ABZ69178;		
Qy	5130	CACCGCAGGCTGATGCTGGAGATGCGCCGAAATGAATCGACAGCTTTTACTTGAGCA	5189	DT	11-AUG-2003	(first entry)	
Db	5850	GACAGAGGTTTGTGCTGGAGATC---TGATTTGAATTAATCCGTGCTTACTCAATCAT	5906	XX	Vector construct pACYCcarol4	coding sequence.	
Qy	5190	GTAACGCAACCATGCGGTGGCTCGAAGAGTTTCGCCACCGCCGCAAGCTGTTGAT	5249	DE	Biosynthesis;	enzyme; antibacterial; protozoacide; herbicide; isoprenoid;	
Db	5907	GCGGTGAAACGATGCGAGTGGCTCGAAAGTTTTCGACAGCCTCAAAGTTATTGAT	5966	KW	inhibitor; immunostimulant;	gene; ds.	
Qy	5250	GCACGACGCGCGCAGCAGCTGATGTGTATGTGTGTGTCGTCACTCGCATGATGTG	5309	XX	Unidentified.		
Db	5967	GCAAAACCCGCGCAGCGTACTGATGCTCTACGCTGTGTGGCCATTTGTGACGATGT	6026	OS	Synthetic.		
Qy	5310	ATTGATGGGCAACGCTGGGCGAAGCGGCGACGCGATGCGGTGGAAGCGCGGAGCA	5369	XX	DE10201458-A1.		
Db	6027	ATTGACGATCAGACGCTGGGCTTTTCAGGCGCGGACGCTGCTTACAAACGCGCCGAA	6086	XX	17-OCT-2002.		
Qy	5370	CTATGCGACATCTGAATTCGAACCCGCGGCTACAGCGCGCGCACATGATGAA	5429	XX	16-JAN-2002; 2002DE-01001458.		
Db	6087	CGTCTGATGCAACTTGAGATGAACGCGCCAGGCTATGCGAGATCGCAGATGCAGAA	6146	XX	11-APR-2001; 2001DE-01018166.		
Qy	5430	CGGCGTTTATGCGGCTTTCAGGAAGTGGCGATCATTCACGAGCTGCGCAACAACTGGCG	5489	PR	22-JUN-2001; 2001DE-01030236.		
Db	6147	CGGCGTTTTCGCGCTTTCAGGAAGTGGCTATGGCTCATGATATCGCCCGGCTACGCG	6206	PR	09-NOV-2001; 2001DE-01055084.		
Qy	5490	TTTGATCATCTGAAGGCTTCGATGATGATGACGCAACGAAATACGCGAGCTTCGAT	5549	XX	(BACH/) BACHER A.		
Db	6207	TTTGATCATCTGAAGGCTTCGATGATGATGATGATGATGATGATGATGATGATGAT	6266	PA	(ROHD/) ROHDICH F.		
Qy	5550	GACAGCTGCGTTACTGCTATCACGTCGCGGCGTGTGCTGATGATGCGCGCTA	5609	XX	Adam P, Amsingler S, Bacher A, Eisenreich W, Hecht S, Rohdich F;		
Db	6267	GATACGCTGCGCTATGCTATCACGTTGCGGCGTGTGCGGCTGATGATGCGCAATC	6326	PI	WPI; 2003-185556/19.		
Qy	5610	ATGGCGCTGCGGACGAAAGCGTGTGCTGATCACGCTGCGATTTAGGACTGCGCTTCCAG	5669	XX	New proteins involved in isoprenoid biosynthesis, useful in screening for		
Db	6327	ATGGCGCTGCGGATTAACCCAGCTGACCGGCTGTGACTTGGGCTGCAATTCAG	6386	XX	inhibitors, also new intermediates, potential therapeutic agents, nucleic		
Qy	5670	CTCACTAACATTGCGCGACATGATGAGATGCGGAAATGCGTGTCTATCTGCGG	5729	DR	Example 7; Page 48-50; 78pp; German.		
Db	6387	TTGACCAATATTGCTCGCATATTGACAGATGCGCATGCGGCGCGCTTATCTGCGG	6446	XX	The present invention relates to proteins capable of the enzymatic		
Qy	5730	CAATCTGCTCGATCAGCGGGATTAACGCGCGGATACGCTGACCTGCACTGCCGCAATCGT	5789	CC	conversion of 2C-methyl-D-erythritol-2,4-cyclodiphosphate to 1-hydroxy-2-		
Db	6447	GCAAGCTGCTGGAGCATGAAGGCTGAAACAAAGAGAAATTAATGCGGCACCTGAAACCGT	6506	CC	methyl-2-butenyl-4-diphosphate. The proteins are useful in screening a		
Qy	5790	GACGCGCTCGCTCACTGCGCAGCGGTTTGTAGTGGGAGCGGAAACCTTATTAATCACTCG	5849	CC	chemical library for inhibitors of isoprenoid biosynthesis. Cells and		
Db	6507	CAGCGCTGAGCGGATATCCCGCTGTTTGTGTGAGGAGCAGAACCTTACTATTGTCT	6566	CC	organisms that express them are useful for in vivo/in vitro production of		
Qy	5850	GCGCATCCGTTTACCGGTTTACCGCTGCGCTCGGCTGGGCGCATCGCTACGCTCGC	5909	CC	isotopically labeled biosynthetic intermediates and products of the		
Db	6567	GCCACAGCGCGCTGCGAGGTTGCCCTTCGCTTCGCGTTCGCGCAATCGCTACGCGGAG	6626	CC	nevalonate-independent pathway of isoprenoid synthesis. Such		
Qy	5910	GCGCTTATCGCGAATTTGGCTCAAGTTTCAGCAGCGCGGTGTGCAACGCTTGGATTC	5969	CC	intermediates and products are used to screen for genes, enzymes or		
				CC	inhibitors of isoprenoid/terpenoid biosynthesis, potentially useful as		
				CC	antibiotics, antimetabolites and herbicides and as immunostimulators,		
				CC	particularly for treating bacterial infection. Antibodies against the		
				CC	intermediates are useful for detecting pathogens, in body fluids. The		
				CC	present sequence is a coding sequence used in the exemplification of the		
				CC	invention		
				XX	Sequence 8547 BP; 1884 A; 2296 C; 2279 G; 2088 T; 0 U; 0 Other;		
				SQ	Query Match 23.1%; Score 1618.8; DB 8; Length 8547;		
					Best Local Similarity 66.0%; Pred. No. 0;		
					Matches 2375; Conservative 0; Mismatches 1217; Indels 6; Gaps 2;		
				Qy	2490	CAGCAGCAAGTCGTCTGGCGAGCGGACCTGATGCGCAGCAATACGATGATTG	2549



Db 2845 CACCGGTGAGCTGCTTTAAGTGGAGCGGCTATGCAACCGCAATTAATGATCTGATTTCTCG 2904  
Qy 2550 TCGGTGCTGGAATGCGGTAATGCTTGAATTCGCTGCGTCTGGTCAATTTGAGCCACAC 2609  
Db 2905 TGGGGGCTGGAATCGCGAATGCGCTTATCGCCCTGCGTCTTCAGCAGCAGCAACCTGATA 2964  
Qy 2610 TGAATGCTGTTGCTGAGAGCGAATGCGCATCCGGCAGGCAATCATACCTGGTCTGTTTC 2669  
Db 2965 TCGGTATTTGCTTATCGACCGGCACCCAGGCGGGGNAATCATACGTGCTGATTTTC 3024  
Qy 2670 ATCAGCGAATCTCAGCGCGCAACAACTTCGTGCTGCTGCAACCGCTGATTAACGGTGTG 2729  
Db 3025 ACCACGATGATTTGACTGAGAGCCAACTATCGTTGATAGCTCCGCTGGTGTTCATCACT 3084  
Qy 2730 GGTGAGTTATCAGGTGGTTCCTGCGCTGCGCGCAACTGCAACGGGGAATTAATGTT 2789  
Db 3085 GGGCCGACTATCAGGTACGCTTTCCACACGCGCTGTAAGCTGAACAGCGGCTACTTTT 3144  
Qy 2790 CCATCGCATCAGCGCAATTTTGCCGCCATCTTTACGCGCGATGGGTGACGATCTGTGA 2849  
Db 3145 GTATTACTTCTCAGCGTTTCGCTGAGGTTTACAGCGACAGTTTGGCCCGCACTTGTGA 3204  
Qy 2850 CAAACACAGCGGTAACACAGGTAAACCCACGAGGTGACGCTGCGGATGCGCTGAAC 2909  
Db 3205 TGGATACCGCGTCCGAGAGGTTAATGCGGAATCTGTTGCGTTGAAAAGGGTCAAGTTA 3264  
Qy 2910 TTGCTGCGCAAGTGTGATGATGTCGGGCTGCGCGCAACGCGCATCTGCAAGTGG 2969  
Db 3265 TCGGTGCGCGCGGTGATGACGGCGGGGTTATGCGGCAATTCAGCACTGACGCTGG 3324  
Qy 2970 GTTATCAGGTGTTCTTGACAAAGTGGCAGCTGCGCAGCGCGCAACCGCTGCAAGCAGC 3029  
Db 3325 GCTTCCAGCGTTTATGSCCAGGAATGSCGATTGAGCACCGCATGGTTTATGCTCTC 3384  
Qy 3030 CGATCTGATGATGACCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 3089  
Db 3385 CCATATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3444  
Qy 3090 CGCTCAGCGCGATCGGCTATTGATTGAAGATACCATTAACGATGATGATGATGATGATG 3149  
Db 3445 CGCTCTGCGGACGAGTTGTTAATGAAGACAGCACTAATTAATGAATGAATGAATGAATGA 3504  
Qy 3150 CGGAGAACCGCTGCTGACGACATGCGCGACTATGCCAATCAGCAAGCTGCAAGCTGA 3209  
Db 3505 ATCCCTGAATGCGCGGCAAAATATTGCGACTATGCGCGCAACAGGGTTGCGAGCTTC 3564  
Qy 3210 GTACGCTGCTGCTGAGAGCAGGCAATTAACGATTAACCTGAGCGGCAACATCGATC 3269  
Db 3565 AGACACTGCTGCGAAGAACAGGGCGCTTACCCCAATTAATCTGTCGGCAATGCGGAG 3624  
Qy 3270 GATTTGCGCAACAGCAGCGCGCAAGCGTGCAGCGGCTGCGCGCGGCTGTTTCATG 3329  
Db 3625 CATTCGCG -- CAGCAGCGCGGCTGCTGATGATGATGATGATGATGATGATGATGATG 3681  
Qy 3330 CCAACACCGGTTACTCTTCCGCTGCGCGTGGCGCTAGCGAGTTGGTAGCGGCTGT 3389  
Db 3682 CTACACCGGCTATTCACTGCGGCTGGCGGTTGCGGCGGCGGCGGCTGAGTGCATTTG 3741  
Qy 3390 TGGCCACCGATGCCCTCAGCTCAGCAACATATGCAACGCTTTCGCGCTGAGTGCATTTG 3449  
Db 3742 ATGTCTTTTACGTGCGGCTCAATTCACCATGCGCAATTAACGATTTTGGCCGCGGCTG 3801  
Qy 3450 GCGAACAGCGATTTTTCGCTGCTAAACCGCATGCTGTTTTCGCGGCTGAGCGCGCAGC 3509  
Db 3802 AGCAGCAGGGCTTTTTCGCAATGCTGAATCGCATGCTGTTTTCGCGGCGCGCGGATTT 3861  
Qy 3510 AGCGCTGGCGGCTGATGCAACGTTTTCGCGCTGATGCGGCTGATGCGGCTTAAATGCGCTTTT 3569  
Db 3862 CACGCTGGCGGTTATGCAAGCTTTTATGCTTTTACCTGAAGATTTAAATGCGCGGTTT 3921  
Qy 3570 ACGCCGCGCACTGCGGCTGCGGATAAACGCGGATTTCTGCGCGCAAGCGCGGCTGC 3629  
Db 3922 ATCGGGAAACACTCAGCTGACCGCTGCGCTAGTATTTCTGAGCGGCAAGCGCGCTGTTTC 3981

Qy 3630 CCATCGGTGAAGCGCTGCGCGCTGTTGAATTTCTGCGAACCAGGAGGAGGAGGAGGAGG 3689  
Db 3982 CGGTATTAGCAGCATTCGAAGCCATTAATGACGACTCATCGTTAAAGAGCGACTACATGAA 4041  
Qy 3690 ACGCACTTATGTANTGCGCAGGCTTTGCGGCGCTGCGCTGCGGATTCGCCCTGCAAGC 3749  
Db 4042 ACCAACTACGCTAATTTGTCAGGCTTCGTTGCGGCTGCGCACTGGCAATTCGTCTACAAGC 4101  
Qy 3750 GCGGGCATACCAACCACTTACTCGAGCAGCGGACAAACCGGGCGGACGCGCTATGT 3809  
Db 4102 TCGGGGATCCCGCTCTTACTGCTTGAACAGTGATTAACCCGCGCGTGGGCTTATGT 4161  
Qy 3810 GTTTGAGCAGCAGTGGCTTTTACCTTCGATGCGGACCCACCGTGATCACCGATCCAGCGC 3869  
Db 4162 CTACGAGGATCAGGGTTTACCTTTGATGAGGCGCGACGGTTATCACCGATCCAGTGC 4221  
Qy 3870 CATCGAAGTGTGTTACGCTGCGCAGGAAATCGCTCAGCGATTAACCTGCGAGCTGATGCC 3929  
Db 4222 CATTAAGAACTGTTTGCACCTGCGCAGGAAACAGTTAAAGAGATATGTGCAACTGCTGCC 4281  
Qy 3930 GGTAAACGCTTCTATCGCTGCTGCGAAGATGGCAACAGCTTGAATACGACAATAA 3989  
Db 4282 GGTACGCGGTTTACCGCTGCTGTTGGAGTCAGGAGAGGTCTTAAATACGATAACGA 4341  
Qy 3990 TCAGCGCTGCTGAGCAGCAGATCGCCACAGTTCAATCCGCAAGATGTAGAAGCTATCG 4049  
Db 4342 TCAAAACCGGCTCGAAGCGCAGATTACAGCAGTTTAAATCCCGCGATGTCGAAGGTTATCG 4401  
Qy 4050 TCAATTTCTGCTTATTCAGTGAAGTATTTAGAGAGGTTATCTGAAACTCGCAGCGGT 4109  
Db 4402 TCAGTTTCTGACTATTTCAOCGCGGTTTAAAGAGGCTTATCTAAAGCTCGGTACTGT 4461  
Qy 4110 GCGGTTCTGCGAGTGGTGACATGCTGCGCGTGGCGCGCAGTTGGGACGCTGCAAGC 4169  
Db 4462 CCCTTTTATGCTTCAGAGACATGCTTGGCGCGCACTCACTGGCGAAGTGCAGGC 4521  
Qy 4170 ATGCGCAGCTGCTACAGCATGTTGGGGAATTTTTCAGGACGATCATCTGCGTCAAGC 4229  
Db 4522 ATGAGAGAGCTTTACAGTAAGTTGCGAGTTACATCGAAGATGAACATCTGCGCAGGC 4581  
Qy 4230 GTTTTCTTCCACTCATTTGCTGGGCGGTAATCTTTTCAACGCTCATCTCATATAC 4289  
Db 4582 GTTTTCTTCCACTCGCTGTTGGGCGGCAATCCCTTCGCCACTCATCTCATTTATAC 4641  
Qy 4290 CTAAATTCATGCTGAGCGTGAATGCGGCGTGTGTTTCCGCGCGCGCAACGCGCGC 4349  
Db 4642 GTTGATACACGCGCTGGAGCGTGTGGGCGTCTGGTTTCCGCGTGGCGCAACGCGCGC 4701  
Qy 4350 GCTGCTGAGGCGATGGCGCGACTGTTGAGGACTTGGGCGCGGAGCTGTTACTGAATGC 4409  
Db 4702 ATTAGTTCAGGGGATGATAAAGCTGTTTTCAGGATCTGGGTGGCGAAGTCGTGTTAAACGC 4761  
Qy 4410 CGAAGTGAAGCAGCTGGAACACAGCGCAATCGCAATAGCGCGTTCAGTTAGAGGCGG 4469  
Db 4762 CAGAGTCAAGCATATGGAACACAGGAAACAGATTGAACCGCTGCAATTTAGAGACGG 4821  
Qy 4470 ACGAGCTTCGATGCGCGCTGTGGCTTCAATGCGGCGAGTGTGCTATCTAGCAAA 4529  
Db 4822 TCGCAGGTTCTGACGCAAGCGCTGCGCTCAATGCAATGTGGTTCATACCTATCGGA 4881  
Qy 4530 ACTGCTGCGCAACATCGCTGGCAATGAACGTGCGCATCGCTGAAGCGTGAAGCGAT 4589  
Db 4882 CCTGTTAAGCCAGCACCTTGGCGGTTAAGCAGTCCAAACACTGCAAGCTTAAGCGCAT 4941  
Qy 4590 GAGCAACTCGCTGTTTGTACTTATTTGGCTGATCAGCGCATTAACAGCTCGCGCA 4649  
Db 4942 GAGTAACCTCTGTTTGTGCTCTATTTTGGTTTGAATCACCATCATGATCAGCTCGCGCA 5001  
Qy 4650 CCACACCGCTGTTTGGCGCGCTTATCGTGAGTTGATCGATGAGATTTTCAACAGCAG 4709  
Db 5002 TCACACGCTTGTTCGCGCGCGTTACCGCGAGCTGATGACGAAATTTTAAATCATGA 5061



Qy	4710	CCAGCTCGGCAGACGATTTTTCACCTTTTACCTGTGCACGCGCCCTGTGCAGCAGCGATCCGTGCGT	4769
Db	5062	TGGCCCTCGCAGAGGACTCTCTCACTTTATCTGCACGCGCCCTGTGTCAACGGATTCGTCACT	5121
Qy	4770	GGCACCGCCGGCTGCGGCAGCTTTTATGTGTATTAGCGCGGTGCGCGATCTCGGCACCGC	4829
Db	5122	GGCGCCTGAAGGTTTGGCGCAGTTTACTATGTGTGTGGCGCCGGTGGCGCATTTTAGGCACCGC	5181
Qy	4830	TGCATCGACTGGCCAAACAGGAAGACCGCGCTTCCGCGCATTCGAATTTTTCCTTATCTGGA	4889
Db	5182	GAACCTCGACTGGACGGTTGAGGGGCCAAACACTACGGACCGTATTTTTCGGTACCTTGA	5241
Qy	4890	GCAGCACTATACGCGCGGATTTACGTTCAGCAAAATTAGTGACACACAGAAATGTTTACGCGCTT	4949
Db	5242	GCAGCATTATACGCTCGGCTTTACGGAGTCAGCTGGTTCACGCACCGGATGTTTACGCGCTT	5301
Qy	4950	TGATTTTCGGCAGACGCTGCATGCCCATCAGGCTCGGCGTTCCTGCTGAGAGCGGATTTT	5009
Db	5302	TGATTTTCGGCAGCAGCTTAAATGGCTATCATGGCTCAGCCCTTTTCTGTGAGAGCCGTTCT	5366
Qy	5010	GAGCGAAAGCGCCTGGTTCCGCCCGCATAAACCGCGATGCGCGATATCAGCAATCTCTATCT	5069
Db	5362	TACCCAGAGCGCTGGTTTCGGCGCATTAACCGCGATAAACCAATTACTTAATCTCTACCT	5421
Qy	5070	GGTGGGTGCGGTTACGATCACGGCGGGCGTGC CGGGCGTGATCGGTTTCGGCCAAAGC	5129
Db	5422	GGTCGGCGCAGGCACGCATCCGCGCGAGGCATTCCTGGCGTCACTCGGCTCGGCMAAAGC	5481
Qy	5130	CACGGCCAGGCTGATGCTGAGGATCGGCGGATGAATCGACAGCCTTTTACTTTGAGCAA	5189
Db	5482	GACAGCAGGTTTGATGCTGAGGATC---TGATTTGAATAATCCGTGCTTACTCAATCAT	5538
Qy	5190	GTAACGCMAACCATGGCGGTGGGCTCGAAAGATTTTCGCCACCGCGCCCAAGCTGTTTGAT	5249
Db	5539	CGGTCMAACGATGGCAGTTGGCTCGAAAGTTTTCGACAGCCTCAAAAGTTATTTGAT	5598
Qy	5250	GCACCGAGCGCGCAGCACGCTGATGCTGTATGCTGTGTGTCTCACTCCGATGATGTG	5309
Db	5599	GCAAAAACCGCGCGACGTTACTGATCTCTACGCTGGTGGCGCCATTTGTGACGATGTT	5658
Qy	5310	ATTGATGGGCMAACGCTGGCGGAAGCGGCACGCGAGCATCCGCTCGAAGACGCGCAGGCA	5369
Db	5659	ATTGACATCAGACGCTGGGCTTTTCAGGCCCGCGAGCTCGCTTTACAAAAGCCCGAACAA	5718
Qy	5370	CGTATGCAGCATCTGCAAAATTGAAAACCGCGCGCCCTACAGCGCGCGCACATGGATGAA	5429
Db	5719	CGTCTGATGCAATTGAGATGAAAACCGGCCGCTATGCAGGATCGCAGATGCACGAA	5778
Qy	5430	CCGCGTTTATGGCGTTTTCAGGAAGTGGCGATCATTCACGAGCTGCCGCAACACTGGCG	5489
Db	5779	CCGCGTTTTCGGGCTTTTCAGGAAGTGGCTATGGCTCATGATATCGCCCCGGCTTACGCG	5838
Qy	5490	TTTGATCATCTGGNAGGCTTCGCTATGGATGCAGCAACGAAATTTAGCGGAGGCTTCGAT	5549
Db	5839	TTTGATCATCTGGNAGGCTTCGCGCATGGATGTACGCGAAGCGCAATACAGCCAACTGGAT	5898
Qy	5550	GACACGCTGGTTACTGCTATCAGTCGCGGGCTGGTTCGTTTGTATGATGGCGCGGTA	5609
Db	5899	GATACGCTGGCTATTGCTATCAGCTTGCAGCGGTTGTCCGCTTGATGAGCGCAATC	5958
Qy	5610	ATGGCGTGCAGCAAGCGGTGCTCGATCAGCGCTCGCATTTAGGACTGGCGTTCCAG	5669
Db	5959	ATGGCGTGCAGGATAACGCCACGCTGCACCGCGCCCTGTGACCTTGGGCTGGCATTTTCAG	6018
Qy	5670	CTCACTAAACATTGGCGCGGCATTTGTATAGAGATGCCGAAAATGGTCCGCTGCTATCTGCCG	5729
Db	6019	TTGACCAATATTGCTCCGATATTGTGTGACGATGTGTGCGTGGTGTGAGCGCGCTGTTATCTGCCG	6078
Qy	5730	CAATCTTGCTCGATCAGGCGGATTAACGCGCGATACGCTGACTGTCACCGCAACATCTGT	5789
Db	6079	GCNAGCTGGCTGGNAGCATGAAGTCTGAAACAAGAGAAATTTATGCGGCACCTGAAAACCGT	6138
Qy	5790	GCAGCGCTCGCCTCACTCGGCAGCGGTTTATGTGGCGGAGCGGAACCTTATTTACTCTCG	5849

RESULT 10

ACF67367 41/C

Continuation (42 of 57) of ACF67367 from base 4100001 (Photorhabdus lu

WP Sequence split into 57 fragments LOCUS ACF67367 Accession ACF67367

WP	Fragment Name	Begin	End
WP	ACF67367_00	100001	110000
WP	ACF67367_01	100001	210000
WP	ACF67367_02	200001	310000
WP	ACF67367_03	300001	410000
WP	ACF67367_04	400001	510000
WP	ACF67367_05	500001	610000
WP	ACF67367_06	600001	710000
WP	ACF67367_07	700001	810000
WP	ACF67367_08	800001	910000
WP	ACF67367_09	900001	1010000
WP	ACF67367_10	1000001	1110000
WP	ACF67367_11	1100001	1210000
WP	ACF67367_12	1200001	1310000
WP	ACF67367_13	1300001	1410000
WP	ACF67367_14	1400001	1510000
WP	ACF67367_15	1500001	1610000
WP	ACF67367_16	1600001	1710000
WP	ACF67367_17	1700001	1810000
WP	ACF67367_18	1800001	1910000
WP	ACF67367_19	1900001	2010000
WP	ACF67367_20	2000001	2110000
WP	ACF67367_21	2100001	2210000
WP	ACF67367_22	2200001	2310000
WP	ACF67367_23	2300001	2410000
WP	ACF67367_24	2400001	2510000
WP	ACF67367_25	2500001	2610000
WP	ACF67367_26	2600001	2710000
WP	ACF67367_27	2700001	2810000
WP	ACF67367_28	2800001	2910000
WP	ACF67367_29	2900001	3010000
WP	ACF67367_30	3000001	3110000
WP	ACF67367_31	3100001	3210000
WP	ACF67367_32	3200001	3310000
WP	ACF67367_33	3300001	3410000
WP	ACF67367_34	3400001	3510000
WP	ACF67367_35	3500001	3610000
WP	ACF67367_36	3600001	3710000
WP	ACF67367_37	3700001	3810000
WP	ACF67367_38	3800001	3910000
WP	ACF67367_39	3900001	4010000
WP	ACF67367_40	4000001	4110000
WP	ACF67367_41	4100001	4210000
WP	ACF67367_42	4200001	4310000
WP	ACF67367_43	4300001	4410000
WP	ACF67367_44	4400001	4510000
WP	ACF67367_45	4500001	4610000
WP	ACF67367_46	4600001	4710000
WP	ACF67367_47	4700001	4810000

WP	ACF67367_48	4800001	4910000	
WP	ACF67367_49	4900001	5010000	
WP	ACF67367_50	5000001	5110000	
WP	ACF67367_51	5100001	5210000	
WP	ACF67367_52	5200001	5310000	
WP	ACF67367_53	5300001	5410000	
WP	ACF67367_54	5400001	5510000	
WP	ACF67367_55	5500001	5610000	
WP	ACF67367_56	5600001	5648894	
Query Match				19.6%; Score 1374.8; DB 10; Length 110000;
Best Local Similarity				62.0%; Pred. No. 0;
Matches 2211; Conservative				0; Mismatches 1347; Indels 8; Gaps 2;
Qy	2532	AATACATGTGATTTTGGTTCGGTCTGACGTCGGCAATGCGCTTCATTTGCGCTCGGCTTCG	2591	
Db	23257	ACTGGGATCTGATTTCTGGTGGAGAGGATTTGGCCAAACGGACTGATCGCAATGGTTTCC	23198	
Qy	2592	GTCAATTTGACGCCAACATCGAATGCTTGTCTCGAGAGCGATCGCATCCGGCAGGCA	2651	
Db	23197	AGCAGTGTAAACACATCTGCGAGTGTGCTTATTTGAAACACAGAAACAATAGAGGCA	23138	
Qy	2652	ATCATACCTGGTGGTTTCATCAACGCAATCTCAGCGCCGAACAACATTCGTCGCTGCAAC	2711	
Db	23137	ATCACACGTGGTCATTTCTATCAACATGATCTTACTTGAGCGCGAAACATGAGTGAATGCAC	23078	
Qy	2712	CGCTGATTACCGTGCCTGGTCAGGTTATCAGTGCCTTTTCTCGCTCGCGCCGAATC	2771	
Db	23077	CGCTGATTATCCTATCCCTGGTCAGTTAGACGTCAATTTTTCAGCATTTCAACGCACAT	23018	
Qy	2772	TGGACGGGGATTAATGTTCCATCGCATCAGGCGAATTTTCCCGCGCACTTTTACCGGGCA	2831	
Db	23017	TGCCACATTCATATTTTCAGTATCACATCCAAACACTTTTGCAGCATATCTCCATGCATATT	22958	
Qy	2832	TGGGTGACGATCTGTGGCAACAACAGCCGTACACAGGTAAACCCACGACGTCGACGC	2891	
Db	22957	TGGCGCAACGTAACACAGCTCGTTTATTTGGTACAGAGCTGACTCCACAGAAATTTACT	22898	
Qy	2892	TGGCGGATGCGCGTGAACCTTGCTCGCAAGTGGTGATTTGATGGTCGCGCCCTGCAGCGGA	2951	
Db	22897	TACAGGACGCTCGTCTCTAAGTGCTGGCGAGTCATTGATGGAGAGGCTGGCGCACAG	22838	
Qy	2952	CGCCACATCTGACGTGGGTTATCAGGTGTTTCTGGACAAGTGGCAGCTGGCGGACG	3011	
Db	22837	GACCATTTATAGGAGTGGCACCCACGCGCAATTTTGGTCAGAAATGGGAGCTGGAAGAT	22778	
Qy	3012	CGCAGCGCTGCAGCAGCCGATCTGATGGATGCCACCGTCGATCAGACACATCGCGACTATGCCAATC	3071	
Db	22777	CGCACTCTTTAAACCCACCGATTTTAAATGGATACTAGTGTGGGACAGGATACAGGTTATC	22718	
Qy	3072	GTTTTGTCTAACGCTGCGCTCAGCGCGGATCGGCTATTTGATTTGAAGATACCAATTACG	3131	
Db	22717	GATTTATCTATGTCCTCGCTTCTCTCAACTCGTCTGTGATAGAGGACACTCATTTACG	22658	
Qy	3132	TTAACCCAGCCGCGCTGGCGGGAACAACCGCTCGTCAGCACATCGCGACTATGCCAATC	3191	
Db	22657	TTGATCGGGGCGCACTTGATAAGGCTTTGTCCGAGCTACTATCCGAGTACGCCGAAGA	22598	
Qy	3192	AGCAAGGCTGGACGCTCAGTACGCTGCTGGTGAAGACGACACGGCATATTACCGATTACCC	3251	
Db	22597	AACATGGATGGAACCTGGTTAAACTCATTCGAAAGAGAGCGGTTGTTCTTCCAAATTACAC	22538	
Qy	3252	TGAGGGGCAACATCGATCGATTCTTGGCAACAGACGCGGGCCAGGCTGCAGGGCCTGC	3311	
Db	22537	TTACGGGAGATTTTAACTCTTTCTGGGCACAGCTAGCAGGACAGCCCACTGTGGGTTAC	22478	
Qy	3312	CGCGCGGCTGTTTCATGCCAACACCGGTTACTCTTTGGCGTCCGCGCTGGGCGCTAGCGG	3371	
Db	22477	GTGCAGCTTGTGTTTCAACCCCAACACAGGCTACTCTCTGCCACACGCTATTTCGTTGGCAG	22418	
Qy	3372	AGTTGGTAGCAGCGTGTGGCCACCGATGCCCTCAGCTCAGCCAAACATATCGAACGCT	3431	
Db	22417	ATCGTATTTGTTGCTCTCGCGGAGCTTACCGATACCTCTTATTTCATTTACCTTACGGGATT	22358	

QY 4507 GAGTGGTGCATACCTACGACAACTGCTTGGCCACCATCCGCTGGCAATGAACGTGGC 4566  
Db 21277 GATGTTATTCATACCTTAAGACACCTGTTAGTGCAGCATCCGCTCGGAATAGCCCGCCT 21218  
QY 4567 ACATCGCTGAAGCGTAAGCGCATGAGCAACTCGCTGTTGTACTCTATATTTGGCTGTGAAT 4626  
Db 21217 CGAACACTGGTGGTAAGCAATGAGCAACTCACTGTTGCTGCTTTATTTTCGGTCTGAAC 21158  
QY 4627 CAGCGCGATGAACAGCTCGCGCACCAACACCGTCTGTTTGGCCCGCTTATCGTAGTTG 4686  
Db 21157 CATCACCATACGCAATAGCCCATCATACAGTTTGTCTTGGTCCAGCTATAAGGAATC 21098  
QY 4687 ATCGATGAGATTTCAACAGCAGCGCATGAGCGATGTTTCACTTTACCTGACCGG 4746  
Db 21097 ATCGAAGATATTTTATCATGCCAGCTGTGAGAGGATTTTCACTCTATCTCATGCC 21038  
QY 4747 CCCTGCAGCAGGATCCGCTCGCTGGCACCGCCCGCTGGCAGAGCTTTTATGTTTAGCG 4806  
Db 21037 CCCTCTGTTACTGACCCATCCCTTGTCTCCCGAGGATGCGAAGTTACTAGCTTCTGGCC 20978  
QY 4807 CCGGTGCCGATCTCGGCAACCGCTGACATCGATGCAACAGGAAGGACCGCGCTTGGC 4866  
Db 20977 CCTGTACCAATTTGGGACAGCTAATTTAACTGGGATATAGAGGGGCCACGCTTACGT 20918  
QY 4867 GATCGAATTTTCTTATCTGAGCAGCACTACATCGCGGATTAGTCAAGCAATTAGTG 4926  
Db 20917 GACCGTATTTTCGCGTATCTGAAAGTACTATATGCCCGGTCTGCTCAAGCAGTTAGTC 20858  
QY 4927 ACACACAGATGTTTACGCGCTTGTGATTTTCGCGACACGCTGATGCCCATCACGCTCG 4986  
Db 20857 GTTCATCGTATTTTACCCATTTGATTTTCTGTGACCACTCAATGCCCATCTCGGTTGG 20798  
QY 4987 GCGTTCCTGAGCGGATTTGACGAAAGCGCTGTTTCCGCCCGCATAACCGCGAT 5046  
Db 20797 GCCTTCCTTTGAACCGTTGCTGACACAAAGCGCTTGGTCCGACCAACAATCGGAC 20738  
QY 5047 GCGGATATCAGCAATCTCTATCTGGTGGTGGCTAGCCATCCAGGCGCGGCTGCC 5106  
Db 20737 AACCGCATTTGATATCTGATCTCGTGGCGGAGGACGATCTCTGGAGCAGGTATCCC 20678  
QY 5107 GCGGTGATCGGTTCCGCCAAGGCCACCGCAGGCTGATCTCGAGGATCGCCCAATGA 5166  
Db 20677 GCGGTGATTTGATCAGCAAAAGCCACCGCCCACTTGATGTTTAGAGGATATCGCTAATGA 20618  
QY 5167 ATGACAGCTTTTACTTTGAGCAAGTAACCAACCATGCGGTGGCTCGAAGATTTGG 5226  
Db 20617 ATCTGCGC-...TACTTAAACAGTTACTCAGATAATGAGCAAGGCTCAAAAAGTTTG 20561  
QY 5227 CCACCGCCGCAAGCTGTTGATGCACGACGCGCGGACGAGCTGATGCTATGCT 5286  
Db 20560 CCAGTGTACCCGACTTTTTCGATACAGCAACACGACACGCAATGATGCTGTACGCT 20501  
QY 5287 GGTGTGCTCACTGCGATGATGATTGATGGGCAACGCTGGCGAAGCGCGACGACG 5346  
Db 20500 GGTGCGGTTATTTGATGATATATAGACGGGCAAGAGCTGGGAGACAAATAAGCAGTG 20441  
QY 5347 ATGCGGTCGAAGACGCGCAGGACGATPACGACATCTGCAATTTGAAACCCCGCGCT 5406  
Db 20440 TTGATAAGTATAGCGCCCGGAGAACTTCAAAATGCTGCAATACCTGACAAAGCAGGCTT 20381  
QY 5407 ACAGCGCGCGCACATGAGTAACCGCGTTTATAGCGGTTTACGAGAGTGGCGATCATTC 5466  
Db 20380 ACAGCGGTTTGGCGATGACCGCAACCTGCTTTTGCAGCCTTTTTCAGACAGTACTGATGA 20321  
QY 5467 ACCAGCTGCGCAACAACTGGGCTTTGATCATCTGGAAGGCTTTCGCTATGATGACGCA 5526  
Db 20320 ATGAAATCCTCAGCAACAGGCTTCGAACATCTGGAAGGATTCGCAATGGATGACTCT 20261  
QY 5527 ACGAACATTCAGCAGCTTCGATGACACGCTCGGTTACTGCTATCACTGCGCGGCTGG 5586  
Db 20260 GTGAGCCTTACAGAACGTTGGATGATCTCTGAGATATTGCTTACCATGTCCGCGGAGTAG 20201  
QY 5587 TCGGTTTGTATGATGGCGCGCTAAATGGGCGTGGCGACGAAGCGGTGCTCGATCAGCCT 5646

Db 20200 TTGGGCTGATGATGCAAGGGTTATGCGGGTTCGCGAGCGAGTGTGCTGGATCGGGCT 20141  
QY 5647 GCGATTTAGGACTGGCGTTCAGCTCACTAACTTGGCGCGACATTTAGTAAGATGCGG 5706  
Db 20140 GCGATCTAGGAATGCTTTTTCAGTTAACTAACATTTGCTCGGATATTTAGGATGCA 20081  
QY 5707 AAAATGGTGGCTGCTATCTGCGCAATCTGCGCTCGATCAGCGCGGATTAACGCGCGATA 5766  
Db 20080 AGGCGGGGCTGCTATCTTCCGCTAGAAATGGCTTTCACCAAGAGGGCTTAATGCCAGACA 20021  
QY 5767 CGCTGACATGCAACCGCAACATGTCGAGCGCTCGCTCACTGCGAGCGGCTTTAGTGGCG 5826  
Db 20020 CACTTATTTATACCAAAACCGCTCTCGCTAGCTCGGGTGGCATCCCGATTTGTTGG 19961  
QY 5827 AGCGGAAACCTTATTTATCACTCGGCGCATCCGCTTTACCGGGTTTACCGCTCGCTCG 5886  
Db 19960 AAGCGGATCTTATTTATACCTCCGCTCGACAGGCTTGTGTTTACCTTTGCGTTGAG 19901  
QY 5887 CGTGGCCCATCGCTACGCTCGCGCTTTATCGGAAATTTGGCGTCAAAAGTTCAAGCAG 5946  
Db 19900 CTTGGCTCATCGCTTCTGCTCATGTTATCTATCTGTAATCGCATTAAGTGCAACAG 19841  
QY 5947 CCGGTGTCACGCTCGGATTCACGCGAGCGCACCAAGTAAAGTGAAGAACTGGCGCTGC 6006  
Db 19840 CAGGAGTACGGCTTGGGATTCAGACAAAGAACTAATCGTGGAGAGAGTTACTTAT 19781  
QY 6007 TGGTGAAGGGCAGGTTTGGCGATCACTTCGCTGTGTCTGCTCTGAAACCGCTCGG 6066  
Db 19780 TAATGGCTGGCATTAAGCGCTTATATCTCGGTGGCGATAACAAGCCTCGCGATC 19721  
QY 6067 CTGCTGTGGCAGCGCTCTCTGTTGA 6092  
Db 19720 CCAATTTGGCAGCGACCTCATTA 19695

RESULT 11  
ACF65388\_06  
Continuation (7 of 13) of ACF65388 from base 600001 (Photorhabdus luminescens nucleoid  
WP Sequence split into 13 fragments LOCUS ACF65388 Accession ACF65388  
WP Fragment Name Begin End  
WP ACF65388\_00 1 110000  
WP ACF65388\_01 100001 210000  
WP ACF65388\_02 200001 310000  
WP ACF65388\_03 300001 410000  
WP ACF65388\_04 400001 510000  
WP ACF65388\_05 500001 610000  
WP ACF65388\_06 600001 710000  
WP ACF65388\_07 700001 810000  
WP ACF65388\_08 800001 910000  
WP ACF65388\_09 900001 1010000  
WP ACF65388\_10 1000001 1110000  
WP ACF65388\_11 1100001 1210000  
WP ACF65388\_12 1200001 1225559

Query Match 19.6%; Score 1374.8; DB 10; Length 110000;  
Best Local Similarity 62.0%; Pred. No. 0;  
Matches 2211; Conservative 0; Mismatches 1347; Indels 8; Gaps 2;  
QY 2532 AATACGATGTGATTTTGGTTCGCTGGACTGGCGCAATGGCTTTGATTCGCTGCGCTCGC 2591  
Db 101680 ACTGGGATCTGATTTCTGTTGGAGGAGGATTTGGCCAAACGGAATGATCGCAATGCGTTTCC 101739  
QY 2592 GTCAATTCGAGCCACAACTGAAATCCCTGTTGCTGGAGAGCGATGCGCATCCGCGAGGCA 2651  
Db 101740 AGCAGTGAACCAACATCTGCGAGTGTTGCTTATTGAAACACAGAAACAATAGAGGCA 101799  
QY 2652 ATCACTGCTGCTGTTTCAACACGATCTCAGCGCGCAACAATTCGCTGCTGCTGCAAC 2711  
Db 101800 ATCACACGCTGCTCAATTTCAATCAATGATCTTACTGAGCGGAAACATGATGATGATGAC 101859  
QY 2712 CGCTGATTTACCGTGGTGGTGGCTGATCAGTGGCTTTTCTCGCTGCGCGCGCAATC 2771

Db	101860	CGCTGATTACCTATCGCTGGTCAGGTTACGAGTCAATTTTTCAGCAATTTCAACGCACAT	101919	Qy	3847	ACGGTGATCACCGATCCCAGCGCCATCGAAGAGTTGTTTACGCTGGCAGGAAAAATCGCTC	3906
Qy	2772	TGGACGGGATTAATTTTCCATCGCATCAGGCGAATTTTTCGCCGCATCTTTTACCGCGGA	2831	Db	103000	ACCGTAATCACCGCTCTTAATGTCTATTGAAGAAATTTTATCCAGGCTGGTAAACGTATG	103059
Db	101920	TGCCACATTCATATTTTCAGTATCAATCCCAACACTTTTGCAAGCATACTCCATGCATATT	101979	Qy	3907	AGCGATTACGTGAGCTGATGCGGTAACGCCCTTCTATCGCTGTGTGCGGAAGATGCG	3966
Qy	2832	TGGGTGACGATCTGTGGCAAAACAACAGCGGTACACAGGTAAACCCACGCGAGTGC	2891	Db	103060	GCTGATTACGTGATTTACTTCCCGTACACCCCTTTTATCGGCTCTGTGGGAGTCGGGT	103119
Db	101980	TGGCGCAACGTATACAGACTCGTTTATTTGGTACAGGAGCTGACTCCACAGAAAGTTTACT	102039	Qy	3967	AAACAGCTTGAATTTACGACAAATATCAGCGCTGCTGGAGCAGCAGATCCACAGTTCAAT	4026
Qy	2892	TGGCGGATGCGCGTGAACCTTGCTGCGCAAGTGGTGAATGATGCTGCGGCCCTGACGCGA	2951	Db	103120	AGATGTTTGAATTTACGACAAATGATCAGCAACATCTGGAAGCACAATTCATACGTTCAAT	103179
Db	102040	TACAGACCGCTCGCTCTTAAGTGTGTGGCGAGTCAITGATGGGAGAGCTGGCGACAG	102099	Qy	4027	CCGCAAGATGTAGAAGGCTATCGTCAATTTTCTGCTATTCTACGTGAAGTATTTAGAGAG	4086
Qy	2952	CGCCACATCTGACGCTGGGTTATCAGGTGTTCTTTTGGACAAGATGGGAGCTGGCGAGC	3011	Db	103180	CCAAGAGATGTTAAAGGATATCGACCTTTCTAGACTATTCCCGGGAAGCCTTTAATGAA	103239
Db	102100	GACCAITTTATAGGAGTGCCACCCAGGCATTTTGTGGTCAAGAAATGGGAGCTGGAAGAT	102159	Qy	4087	GGTTATCTGAACTCGGCACGCTGCGGCTTCTCAGGTCGCTGACATGCTGCGCGTCGG	4146
Qy	3012	CGACCGCTGACGAGCGGATCCTGATGGATGCCACCGTCCATCAGCAAGCGGTTATC	3071	Db	103240	GGCTACCTGAACTTGGAAACGGTGCCTCTCTTTTCGGGACATGTTAGTGGCGCT	103299
Db	102160	CGCACTCTTTTAAACCCACCGGATTTTAAATGGATAGTGTGGGACAGATACAGTTATC	102219	Qy	4147	CCGAGTTGGGACGCTGCAAGCATGGCGCAGCGCTCTACAGCATGCTGCGGAAATTTAT	4206
Qy	3072	GTTTTGTCTACACGCTGCGGCTCAGCGCGATCGGCTATTTGATGAAGATACCCATTAG	3131	Db	103300	CCACAGTTGATACGTTCTACAGCATGGCGTAGTGTATTATAGTCAGGTGCGCACGCTTATC	103359
Db	102220	GATTTATCTATGCTCTGCGGTTCTCTCACTCGTCTGATAGAGACACTCATTTAG	102279	Qy	4207	CAGSACGATCATCTGCGTCAAGCGTTCCTTCACCTCATTTGCTGCTGCTGGCGGTAATCCT	4266
Qy	3132	TTAACACGCGCGCTGGGAGGAAACACCGCTCGTCGTCAGCACATCCGCGACTATGCCAATC	3191	Db	103360	AAGGATGAGAGCTTGGCCAGGCAATTTTCAATTTTCACTCGCTATTGATAGGAGGTAAACCC	103419
Db	102280	TTGATCGGGGCGCACCTGATGAAGGCTTTGTGCGAGGCTACTATCGCAGAGTAGCGAAGA	102339	Qy	4267	TTTTCAACGTCATGATCTATACCTTAATTCATGCGCTGGAGCGTGAATGGGCGTGTGG	4326
Qy	3192	AGCAAGGCTGGAGCGCTGAGTACGCTGCTGCTGAAGACACCGCATATTTACCGATTACCC	3251	Db	103420	TTTCTCGATCTTCTATCTATATCTTTGATCCAAGCATGCGTAGTGTGAGTGGGAGTCTGG	103479
Db	102340	AAATGATGGAATCTGGGTAACTCAITTCGAGAGAGCGGTTGTCTTCCAAATTAAC	102399	Qy	4327	TTTCCGCGCGCGCACCCGCGCTGGTGACGGGANTGGCGGACGCTGTCGAGGACTTG	4386
Qy	3252	TGAGCGCAACATCGATCGAATCTGCGAACAGCAGCGCGGCGAAGGTCGACGGCGCTGC	3311	Db	103480	TTTTCCGCGAGGAGAAACAGCGCTCTGGTTGAGGCAATGGTGAATTTGTTTACTGACAT	103539
Db	102400	TTACGGGAGATTTTACCTCTTTCTGGGCAAGCTAGCAGGACAGCCACCTGTGGGTTAC	102459	Qy	4387	GGCGCGAGCTGTTACTGTAATGCGAAGTGAAGCGAGCTGGAACACGCGCAATTCGCAT	4446
Qy	3312	CGCGCGGCTGTTTCATGCCACACCGGTTACTCTTTGCGCTCGCGCTGGGCTAGCGG	3371	Db	103540	GGTGAGAGAAATGAGCTTTAATGCAAGGTAAACGCTTACCACCCATGGAATCGGGTC	103599
Db	102460	GTGAGCTTTGTTTACCCCAACAGGCTACTCTCTGCCACAGCTATTCTGTTGGCAG	102519	Qy	4447	AGCGCGTTCAATTTAGAGGCGGACGACGCTTCGATGCGCGCTGTGGCTTCAATGCC	4506
Qy	3372	AGTTGTGAGCAGCGCTTTGCCACCGATGCCCTCAGCTCAGCCCAACATATCGAACGCT	3431	Db	103600	ACAGGTGTCAAATTTAGCTGATGGAGCAATATGACATGTGACGTTGTGGCTCAATGCC	103659
Db	102520	ATCGTATTGTTGCTCTCGGAGCTTACCGATACCTCTTATTCAITTAACCTCAGGAT	102579	Qy	4507	GACGTGTGATACCTACGACAACTGCTCGCCACCATCGCTGGCAATGAACGTCGCG	4566
Qy	3432	TTGCCGCTGACAGTGGCGGGAACAGGATTTTCTGCTGCTGCTAAACCGCATCTGTTT	3491	Db	103660	GATGTTATTACATCTATTAAGCACCTGTAGTCAGCATCCGTCGGATAGCCCGCT	103719
Db	102580	ACGCACGACAAACAGTGGCAACACAGCGCTTTTTCGCTCTTAAATCGCATGCTCTTCC	102639	Qy	4567	ACATCGCTGAAGCGTAAGCGCATGAGCAACTCGCTGTTTGTACTCTATTTTGGCTGAAT	4626
Qy	3492	TGGCGGTAAGCGCAGCAGCGCTGGCGCTGATGCAACGTTTTTACCGGCTCGATGCCG	3551	Db	103720	CGAACACTGGTGGTAAACGATGAGCAACTCACTGTTCTGCTTTTATTTCTGCTGAAC	103779
Db	102640	TCGCTGGGATCCACACAGTGTGGCAGTATGCAACGTTTCTATCAACTTTCCCAA	102699	Qy	4627	CAGCGCATGAACAGCTCGCGACACACCGCTGTTTGGCCCGCTTATCGTAGTTG	4686
Qy	3552	GTTTAATTTAGCGCTTTTACCGCGGCAACGTCGCGCTCGCGATTAACCGGATTTCTGT	3611	Db	103780	CATCACCATACCAATTAAGCCATCATACAGTTTGTCTTGTGTCAGCTTAAGGAATC	103839
Db	102700	ATCTGATTGACAGTTTATCGGAGCAACTTAATTCGTCGACAGGCGCGATTCTCA	102759	Qy	4687	ATCGATGAGATTTTCAACAGCGCAGCTGGCAGACGATTTTTCATCTTACCTGCAACG	4746
Qy	3612	CGCGCAGCGCGGCTGCCATCGTGAAGCGCTGCGCGCTGTGAA-----TTCTGT	3666	Db	103840	ATCGAAGATATTTTTTATCATGACCGACTGTGAGGATTTTTCATCTATCTCCATGCC	103899
Db	102760	TAGGTAAACCAACAGTGGCGGATTAAGGTTGCTTAAGGCAATGTTTAAACACACAAGA	102819	Qy	4747	CCCTGACGACGATCCGTCGCTGGCACCGCCCGCTCGCGAGCTTTTATGTGTAGCG	4806
Qy	3667	CGAACCGAGGAAGAAAATGAACGCACTTATGTGATGGCGGAGCTTTGGCGCTG	3726	Db	103900	CCCTCTGTTTACTGACCCATCCCTTCTCCGAGGATGCGAAGTACTAGCTTCTGCCC	103959
Db	102820	AGCTTCAGGTTTTTATTTATGATTAAGCGCTGGTAAATTTGTTGTTTGTGGCTG	102879	Qy	4807	CGGTCGCGCATCTCGGCACCGCTGACATCGCATCGGCAACAGGAAGGACCGGCTTGGC	4866
Qy	3727	GGCTGGGATTCGCTCGAGCGCGGCAATACCAACCTTTACTCGAGCAGCGGAC	3786	Db	103960	CCTGTACCAATTTGGGACAGCTAATTTAACTGGATATAGAGGGCCAGCTTACGT	104019
Db	102880	GCACTGGCAATTAAGGCTCCAGTCTCGGGGATTCGACATGATTTTGGAGCAACGGAT	102939	Qy	4867	GATCGAATTTTGTCTTATCTGGAGCAGCACTACATGCGCGGATTAAGTCAAGCAATTTAG	4926
Qy	3787	AAACCGGGCGGACGGCGCTTATGTTGAGGACAGTGGCTTTAATCTTGCATGCGGACCC	3846	Db	104020	GACCGTATTTTTCGCTATCTGGAAAGTACTATATGCGCGCTGCTCAAGCAGTTAGTC	104079
Db	102940	AAACAGGTGGACGCTTATGTTTATAAGGAACAGGATTCACCTTTGATCGCGGCC	102999				





Db 2280 TTGATCGGGGCCACCTGATAGGCTTTGTCGAGGCTACTATCGAGATACGCGAAGA 2339  
QY 3192 ACAGAGGCTGAGCGCTGATAGCTGCTGCGTGAAGACGACGGCATATTACGGATTACCC 3251  
Db 2340 AACATGGATGGAACATGGGTAACTCATTCGAGAAAGAGCGGTGTCTTCCAAATTACAC 2399  
QY 3252 TCAGCGGCAACATCGATGATTCTCGCAACAGCAGCGCGGCGCAAGCGTGCAGCGCCCTGC 3311  
Db 2400 TTACGGGAGATTATTACCTCTTCTGCGCACAGCTAGCAGGACAGCCACCTGTGGGTAC 2459  
QY 3312 GCGCGGGCTGTTTCATGCCACCAACCGGTTACTCTTGCCTGCGTGCCTGCGGCTGAGCGG 3371  
Db 2460 GTGAGCTTTGTTTACCCCAACAGAGGCTACTCTCTGCCACACGCTATTGCGTTGGCAG 2519  
QY 3372 AGTTGGTAGCAGCGCTGTGCCCCACCGATGCCCTCAGCTCAGCCAAACATATCGAACGCT 3431  
Db 2520 ATCGTATTGTTGCTCTGCGGAGCTTACCGATACCTCTCTTATTATTAATACCTCAGGAT 2579  
QY 3432 TTGCGCGCTCAGCAGTGGCGGCAACAGCGATTTTTCGCTGCTGCTAAACCGCATGCTGTTTT 3491  
Db 2580 ACGCAGCAACACAGTGGCAACACAGCGCTTTTTCGCTCTTCTAAATCGCATGCTCTTCC 2639  
QY 3492 TGCGCGGTAAAGCGCAGCAGCGCTGGCGGTGATGCAAGTTTTACCGGCTCGATGCG 3551  
Db 2640 TCGCTGGGATCCACAAACGCTGGCAGGTAATGCAACGTTTCTATCAACTTTCCCAA 2699  
QY 3552 GGTAAATTAGCGCTTTTACGCGGGCAACTGCGCTGCGGATAAACCGCGATTCTGT 3611  
Db 2700 ATCTGATTGCAAGTTTTATGCGGAGCAACTTAATTCGTGCACAGGCCGGAATCTCA 2759  
QY 3612 GCGGCAAGCGCGGCTGCCATCGGTGAAGCGCTGCGGCGCTGTGAA-----TTCGT 3666  
Db 2760 TAGGTAACACACAGTGGCGGATAAGGGTGCTCTAAAGGCAATGTTTAAACAAACAAGA 2819  
QY 3667 CGAACAGGAGAAATAAGAAACGCACTTATGATGCGCGCAGGCTTTGCGCGCTG 3726  
Db 2820 AGCTTCAGGGTTTTTATTATGATTAAGCGCTGGTAATTTGGTGTGCTGTTTTGTTGGGCTG 2879  
QY 3727 GCGCTGGCATTCGCTGCAAGCGGCGGCATACCAACACCTTACTCGAGCAGCGCGAC 3786  
Db 2880 GCATGGCAATAGGCTCCAGTCTGCGGGGATCCGACATGATTTTGGAGCAACGGAT 2939  
QY 3787 AAACCGGGGAGCGCGCTATGTTTGAAGCAGGTGGCTTTACCTTCGATGCGGACCC 3846  
Db 2940 AAACAGGTGGACGCGCTATGTTTATAAGGAACAGGGATTCACCTTTGATGCGGCGCC 2999  
QY 3847 ACGGTGATCACCGATCCAGCGCATCGAAGAGTCTTCACGCTGGCAGGAAATCGCTC 3906  
Db 3000 ACCGTAATCACCGCTCTTAATGTCATTGAAGAAATGTTTACCAGCGCTGTAACGATG 3059  
QY 3907 AGCGATTACGTCGAGCTGATGCGGTAACGCCCTCTATCGCTGTGCTGGGAAGATGCG 3966  
Db 3060 GCTGATTACGCTGATTTACTTCCGTAACCCCTTTTATCGGCTCTGTTGGGAGTCGGGT 3119  
QY 3967 AAACAGCTTGAATACGCAATATACCGCTGCTCGAGCAGCAGATCGCCAGTTCAAT 4026  
Db 3120 AAGATGTTTGAATACGCAATATGATGCAACATCTGGAAGCAAAATCCATAGTTCAAT 3179  
QY 4027 CCGCAAGATGTAAGGCTATCGTCAATTTCTTGGCTATTTCAGCTGAAGTATTAGAGAG 4086  
Db 3180 CCAAGAGATGTTAACGGGTATCGACGTTTCTAGACTATTTCCCGGAGACCTTTAATGAA 3239  
QY 4087 GGTATCTGAAACTCGGCAAGGCTGCGGTTTCTCGAGGTGGGTGACATGCTGCGCTGCGG 4146  
Db 3240 GGCTACCTGAAACTTGAACGGTGCCCTTCTCTCTTTTCGCGACATGTTGAGTGGCGCT 3299  
QY 4147 CCGCAGTTGGAGCTGCTGAAGCATGCGCAGCGTCTACAGCATGTTGGCGAAATTTATT 4206  
Db 3300 CCACAGTTGATACGCTTACACGCAATGGCGTAGTGTATTAGTCAGTTGGCACCGCTTATC 3359  
QY 4207 CAGGACGATCATCTGCGCTCAGGCGTTTTCTTCCACTCATTTGCTGGTGGCGGTAATCT 4266  
Db 3360 AAGGATGAGAGCCTGGCCAGGCATTTTTCATTTCACTCGCTATTGATAGGAGTAACCC 3419

QY 4267 TTTGCAACGTCACTGATCTATACCTTAATTATGTCGCTGAGCGTGAATGGGCGGTGG 4326  
Db 3420 TTTGCTGCACTTCTATCTATATCTTTGATCCACGCACTGGAACGAGATGGGAGTCTGG 3479  
QY 4327 TTTCCGCGCGCGCGCACCGCGCGCTGCTGTCAGGGCATGCGCGACTGTTTCGAGGACTTG 4386  
Db 3480 TTTCCGCGAGGAGAACACAGCGCTCTGGTTGAGGCAATGGTGAATTTGTTTACTGACATT 3539  
QY 4387 GCGGCGAGCTGTTACTGAAATGCGAAGTGCAGCTGGAACACGAGCGCAATCGCAT 4446  
Db 3540 GGTGGAGAAATGAGCTTAATGCAAGTAAACGCTTACACACCCATGCAATCGGGTC 3599  
QY 4447 AGCGCGCTTCAGTTAGAGGCGGACGACGCTTCGATCCCGCGCTGFGGCTCCAATGCC 4506  
Db 3600 ACAGTGTCCAAATTAGCTGATGACGAAATATGACATGTGACGTTGTGGCTCAAATGCC 3659  
QY 4507 GACGTGTGATACCTACGACAACTGCTTTCGCCACCATCCGCTGGCAAGAAACGTCG 4566  
Db 3660 GATGTTTATCATCTATAAGCACCTGTAGGTGAGCATCCGTCGGAATAGCCCGCT 3719  
QY 4567 ACATCGCTGAAGCGTAAGCGCATGAGCAACTCGCTGTTGTACTCTATTTTGGCCTGAAT 4626  
Db 3720 CGAACACTGGTGTGCTAAACGATGAGCACTCACTGTTGCTGCTTATTTTCGGTCTGAAC 3779  
QY 4627 CAGCGCATGAAGCTCGCGCACCAACCGCTGTTTGGCGCGCTTATCGTGATTG 4686  
Db 3780 CATCACCATACGCAATTAGCCCATCATACAGTTTGTCTTCCAGCTATAAGAACTC 3839  
QY 4687 ATCGATCAGATTTTCAACAGCAGCGACGTGCGACAGATTTTTCATTTACTTCGACGCG 4746  
Db 3840 ATCGAAGATATTTTATATCATGACCGACTGTGAGAGATTTTTCATCTATCTCCATGCC 3899  
QY 4747 CCCTGCGACGCGATCGTCTGGCACCCCGCTGCGGACGCTTTATGTGTTAGCG 4806  
Db 3900 CCCTCTGTTACTGACCCATCCCTGCTCCCGAGGATGCGCAAGTTACTAGTTCTGGCC 3959  
QY 4807 CCGGTGCGGATCTCGGCACCGCTGACATCGACTGGCAACAGGAAGACCGCGCTGCGC 4866  
Db 3960 CCTGTACCATTTTGGGACAGCTAAATTTAACTGGGATATAGAGGGGCCACGTTTACGT 4019  
QY 4867 GATCGAATTTTGTCTTATCTGAGCAGCACTTACATGCGGGGATTCGTCAGCAATTAGTG 4926  
Db 4020 GACGATATTTTCGCTGATCTGGAAGATGCTATATGCTCCGCTGCTCAAGCAGTAGTC 4079  
QY 4927 ACACAGAAATGTTTACCGCGTTTGAATTTTCGCAACGCTGCTGATGCCCATCAAGCTCG 4986  
Db 4080 GTTCATCGTATTTTACCCCAATTTGATTTCTGTGACCAACTCAATGCGCATCTCGGTTG 4139  
QY 4987 GCGTTTTCGCTGGAGCGGATTTTTCAGCAAGCGCTGTTCCGCGCGCATAAACCGCAT 5046  
Db 4140 GCTTCTCTCTTTGAACCGTTGCTGACCAAGCGCTGTTCCGACCAACAATCGGGAC 4199  
QY 5047 GCGGATATCAGCAATCTCTATCTGTTGGTGCCTGATGCGATCCAGGCGCGGGGTCGCC 5106  
Db 4200 AACCGATTGATAATCTGTAICTGCTGCGCGCAGCAGCATCTCTGGAGCAGTATTCCC 4259  
QY 5107 GCGGTGATCGGTCGGCCAAAGCGCACCGCTGATGCTGGAGGATCGCGCGCAATGA 5166  
Db 4260 GCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4319  
QY 5167 ATCGACAGCTTTTACTGAGCAAGTAACGCAAAACCATGCGGTGGGCTCGAAGAGTTTCG 5226  
Db 4320 ATCTGCGG---TACTTAAACAGCTTACTCAGATATGAGCAAGGCTCAAAAAGTTTTG 4376  
QY 5227 CCAACCGCGCCAAAGCTGTTGATGACCGCGCGCGCAGCAGCTGATGCTGATGCTG 5286  
Db 4377 CCAAGTGTCAACCGACTTTTTCGATACAGCAACACGACACAGCAATGATGCTGACGCT 4436  
QY 5287 GGTGCTGCTACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5346  
Db 4437 GGTGCGGTTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4496



Qy	5347	ATGCCGTCGAAGACCGCGAGCGACGTTATGCGAGCATCTGCGAAATCTGAAACCCGCGCGCGCT	5408
Db	4497	TTGATAAGTATAGCGCCCGCGAGAAACTTCAAATGCTGCAATACCTGCACAAAGCAGCGCTT	4556
Qy	5407	ACAGCGCGCGCACATGATGAAACCGCGTCTTTAGGGCGTTTCAGAAAGTGCAGATCATTC	5466
Db	4557	ACGACGGTTTGCCGATGACCGAACCCTGCTTTGCAGCGCTTTCAGACAGTAGCTCTGAGTA	4616
Qy	5467	ACCAGCTGCCGCAACAACTGGCGTTTGATCATCTGGAAGGCTTCGCTATGGATGCACGCA	5526
Db	4617	ATGAAATCCCTCAGCAACAGGCGCTTCGAACATCTGGAAGGATTTCGAATGGAATGACTCT	4676
Qy	5527	ACGAAACATTACGCGAGCTTCGATGACACGCTGCGTTACTGCTATCACTGTCGCGGGCGTGG	5586
Db	4677	GTGAGCCCTTACAGAAAGCTTGGAATGATCTCTGAGATATGCTACCATGTTCGCGGGAGTAG	4736
Qy	5587	TCGGTTTGTGATGCGCGCGCTAATGGCGTGCSCGACGAAGCGTGCTCGATCACGCGCT	5646
Db	4737	TTGGGCTGATGATGGCAAGGTTATGGGGTTTCGCGAGCGAGTGTCTCGATCGGGCGCT	4796
Qy	5647	GCGATTTAGGACTGGCGCTTCAGCTCACTAACTTTGGCGCGCACTTTGTAGAAGATCCGG	5706
Db	4797	CGCATCTAGAAATTGCTTTTCAGTTAACTTAACATTGCTCGGGATATTATTGAGGATCGGA	4856
Qy	5707	AAAATGTCGCTGCTATCTGCCGCAATCTGGCTCGATCAGCGCGGGAATTACGCGCGGATA	5766
Db	4857	AGGCGGGCGCTGTCTATCTTCGCTAGAAATGGCTTCCACGAAGAGGGCTAATGCGCAGACA	4916
Qy	5767	CGCTGACTGCACCGCAACATGTCAGCGCTCGGCTCACTCGGACAGCGGTTTAGTGCGGG	5826
Db	4917	CACTTATTTATACCGAAACCGTCTCGCGCTAGCTCGGGTGGCATCCGGAATTGAATTGTGG	4976
Qy	5827	AGGCGGAACCCCTATTATCACTCGCGCGGATCCGGCTTTACCGGTTTACCCTGCGCTCGG	5886
Db	4977	AAGCGGAATCTTATTATACCTCCGCTCTGACAGGGCTTGTTGGTTTACCTTTGCGTTTTCAG	5036
Qy	5887	CGTGGGCAATCGCTACGGCTCGCGCGTTTATCCGGAATTTGGCGTCAAAAGTTGACGACG	5946
Db	5037	CCTGGGTCATCGCTTCCTGCTCATGGTATCTATCGTGAATCGGCATTTAAAGTGCACAAAG	5096
Qy	5947	CCGGTGTGCACCGCTGGGATTCAGGCGACCGCACCTAAAGGTGAAAACCTGGCGCTGC	6006
Db	5097	CAGGATACGGGCTCGGGAATGCAACAAGAACTAATCTGTGGAGAGAAGGTTACCTTAT	5156
Qy	6007	TGCTGAAAGGGCGAGGTTTGGCGATCACTTCGCGTGTGTCTCGTCTCGAAACCGCGTCCGG	6066
Db	5157	TAATGGCTGGCGCAATGAAGCGCTTATATCTCGGTTGGCGATTAACAACGCTCCGATC	5216
Qy	6067	CTGCTCTGTGGCAGCGTCTCTGTTGA	6092
Db	5217	CCAAATTTGTGGCACCGACCTCATTA	5242

## RESULT 13

AAV40146/C

ID AAV40146 standard; DNA; 8625 BP.

XX

AC AAV40146;

XX  
DT 10-AUG-1999 (first entry)

DE Flavobacterium sp. R1534 nucleotide sequence.

XX  
KW Carotenoid; pigment; canthaxanthin; R1534; crtB; prephytoene synthase;  
KW crtI; phytoene desaturase; crtY; lycopene cyclase; GGPP synthase; crtE;  
KW crtW2396; beta-carotene beta-oxygenase; food product; fermentation; ds;  
XX *Photobacterium* sp. 1220; *Neisseria meningitidis*.

OS *Flavobacterium* sp.

**XXXXXXXXXXXX**

PN JP10155497-A

XX

3086 CTGCCGCTCAGCGCCGATCGGCTATTGATTGAAGATACCCATTACGTTAACCAAGCCCCGCG 3145

Db 6393 CTGCGCTTCAGTCCACCGCATCTGATCAGAGATACCGCTACAGCGACGGCGCGAT 6334  
QY 3146 TTGGCGGAGAACACCGCTCTGTGAGACATCGCGGACTATGCAATCAGCAAGGCTGAGCG 3205  
Db 6333 CTGACGATGGCGCGCTGCGCAGCGCTCGCTGGAATATGCGCGCAGCGCGGCTGGACC 6274  
QY 3206 CTGAGTACGCTCTGCGTGAAGAGACAGCGCATATTTACCGATTTACCTGAGCGGCAACATC 3265  
Db 6273 GGGCAGGAGATGCGCGCGGGAAGGGGCAATCTGCCCATCGCGCTGGCCCATGACGCCATA 6214  
QY 3266 GATCGAATTTCTGCAACAGCAGCGCGGCCAAGCGTGCAGCGCTGCGCGCGGCTGTTTT 3325  
Db 6213 GGTCTTGGCGGACCAACGCGCAGGGGCGGTGCGGTTGGCTGGGGGCAAGGCTGTTC 6154  
QY 3326 CATGCCACACCGGTACTCTTGGCGCTCGCGCTCGCGCTGAGCGGAGTTGGTAGCAGCG 3385  
Db 6153 CACCCCGTCACCGGATATTCGCTGCCCTATGCGCGCGCAGGTCGCGGATGCCATGCGCGCG 6094  
QY 3386 CTGTTGCCACCGATGCCCTACGCTCAGCCAGCAACATATCGAAACGCTTTGCCCGTCAGCAG 3445  
Db 6093 CCGGACCTGACGACCGCGTCCCGCTCGCGCGGTGCGGGCTGGGCCATCGATCGCGCG 6034  
QY 3446 TTGGCGGAACAGCGAATTTTCGCTCTGCTTAAACCGCATGCTGTTTTTGGCGGTTAAGCGG 3505  
Db 6033 GATCGCG--ACCGCTTCTGCGGTGCTGAAACCGGATGCTTCCGCGGCTGCCCGGCC 5977  
QY 3506 CAGCAGCGCTGGCGGTGATGCAAGTTTACCGGCTCGATGCCGGGTAAATAGCGCG 3565  
Db 5976 GACCGTCGATCGCTGCTGAGCGGTTCTACCGCTTCGCGCAGCGCGCTGATCAGCGCG 5917  
QY 3566 TTTTACCGCGGCAACTGGCGCTGCGCGATAAACCGCGAATCTCTGCGGCAAGCGCGCG 3625  
Db 5916 TTCTATGCGCGGCGCTGACATTTGCGGACCGGCTTCGATCGTCAACCGGACCGCGCGC 5857  
QY 3626 GTGCCATCGGTGAAGCGCTGCGCGCTGTTGAATTTCTGTGCAACCGAGGAGAAAAA 3685  
Db 5856 ATTCTGCTGCGAGCGCGCTGCGCTGCGCGAAGCGCGCTGCTGCGAGGAGAGCA 5797  
QY 3686 TGAAAGCACTTATGTGATGCGGAGGCTTTGGCGGCTGCGCTGCGGATTTCCGCTGC 3745  
Db 5796 TGAGTTCCGCGCATGCTATCGCGCGCAGGTTTCGCGGCTTTCGCTTGCATCGCGCTGC 5737  
QY 3746 AAGCGCGGCGCATACCAACACCTTATCTCAGCAGCGCGCAAAACCGGCGGACGCGCT 3805  
Db 5736 AATCGCGCGGATCGCGACCAACCATCGTGAAGCGCGCGCAAGCGCGCGCGCGCT 5677  
QY 3806 ATGTGTTGAGGACAGTGTCTTACCTTCGATGCCGGAACCGCGGTGATCAGCGATCCCA 3865  
Db 5676 ATGTCTGGAACGATCAGGGGCGACGCTTCTGATGCGAGGCGCGCGCTGTCGACCGACCGG 5617  
QY 3866 GCGCCATCGAAGATTTGTTCAAGCTGGCAGGAAATCGCTCAGCGAATACGCTGAGCTGA 3925  
Db 5616 ACGCTGCGAGAGCTGTGGGCGCTCAGCGGCGCAACCGATGAGCGGTGACGCTGC 5557  
QY 3926 TGCGGTAAAGCGCTTCTATCGCTGTGCGGAGATGGCAAAACAGCTGTGATTAAGACA 3985  
Db 5556 TGCGGTCTCGCGCTTCTACCGGCTGACATGGCGGACGCGCGCGAGCTTCGAATACGTA 5497  
QY 3986 ATAAATCAGCGCTGTCTGAGGACGACATCGCCACGTTCAATCGCAAGATGTAGAGGCT 4045  
Db 5496 ACAGACGACGAGCTGATCCGCGAGTCCGCTCTTCAATCCGCGCGATGTCGATGGCT 5437  
QY 4046 ATCGTCAATTTCTTGCTATTACGTTGAAGTATTTAGAGGGTTATCTGAAACTCGCA 4105  
Db 5436 ATCGCGCTTCCAGATTACCGCGAGGAGGTCTATCGGAGGGGTATCTGAAGCTGGGA 5377  
QY 4106 CGGTGCGGTTCTGCAAGGTGCGTGAATGTCGCGCTGCGCGCGGAGTTGGGAGCTGCG 4165  
Db 5376 CCACGCGCTTCTTGAAGCTGGGCGAGATGCTGAACCGCGCGCGGCGCTGATGCGCGCTGC 5317  
QY 4166 AAGCATGGCGAGCTCTACAGATGTGGCGAAATTTATTAGGACGATCATCTCGCTC 4225

Db 5316 AGGCATACCGCTCGGTCCACAGCATGGTGGCGCGCTTCTATCAGBACCCCGCATCTCGCGC 5257  
QY 4226 AGGCTTTTCTCTTCCACTCACTGCTGGTGGCGGTAACTCTTTTTCGAAACGCTCATCTGATCT 4285  
Db 5256 AGGCTTCTCTGTTCCACACGCTGCTGGTGGCGGGAACCCGTTTTCGACCGATCTGATCT 5197  
QY 4286 ATACCTTAATTCATGCGCTGGAGCGTGAATGGGGCGGTGTGGTTTTCGCGCGGCGGACCG 4345  
Db 5196 ATGCGCTGATCCATCGCTGGAACCGCGCGCGCGCTCTGGTTTCGCAAGGGCGGCAACA 5137  
QY 4346 GCGGCTGTGTCAGGGGATCGCGCTGTTTCAGGAGACTTGGGCGGAGCTGTACTGA 4405  
Db 5136 ACCAGTGGTGGCGGCGATGCTCGCCCTGTTTCAGCGCTCTTGGCGGACGCTGCTGCTGA 5077  
QY 4406 ATGCCGAAGTGCAGCACTGGAACACGAGCGGCAATCGCATTTAGCGGCTTCAGTTAGAGG 4465  
Db 5076 ATGCCCGGTCACGCGGATCGACACGAGGGCGATCGCGCAACGCGGCTCACGCTGCTGG 5017  
QY 4466 GCGACGACGCTTCGATGCGCGCGCTGTGCGCTCAATGCGGAGCTGGTGCATACCTACG 4525  
Db 5016 ACGGCGGCGAGTTTGGCGCGGATACGCTGGCGGCAACGCGGCGTGTGATGCACAGCTATC 4957  
QY 4526 ACAGACTGCTTCGCCACCATTCGCTGCAATGAAACGTCGACATCGCTGAAGCGTAAAGC 4585  
Db 4956 GCGACCTGTGCGGCAATACCGCGCGCGGCGCACCAAGGCGCGATCTGNAACGCGCAGC 4897  
QY 4586 GCATGAGCAACTGCTGCTGTTGTAATCTTATTTGCGCTGAAATCAGCCGATGAAACAGCTCG 4645  
Db 4896 GCTGCTGATGCTGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4837  
QY 4646 GCGACCAACGCTGTTTGGCGCGGCTTATCGTGAGTTGATCGATGAGATTTTCAACA 4705  
Db 4836 CCGACCAACGAGCTTCTTGGCGCGGCTCAAGGGGCTGGTGAACGAGATCTTCAACG 4777  
QY 4706 GCAGCAGCTGGCAGACGATTTTCTTACTTACCTGCAACGCGGCTTGCAGCAGCGATCCGT 4765  
Db 4776 GCGACGCTGCGGACGATTTCTCGATGATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4717  
QY 4766 CGCTGCAACCGCGCTGCGGCGAGCTTTTATGTTGTTAGCGCGGCTGCGGATCTCGGCA 4825  
Db 4716 GCCTGGCGCGCGAGGGGATGTCACGATTAAGCTTCTTGGCGCGCTTCCGATCTGGGCGC 4657  
QY 4826 CCGCTGACATCGACTGGCAACAGGAAGACCGCGCTTGGCGGATCGAATTTTGTCTATC 4885  
Db 4656 GCGCGATGTCGATTTGGAGCCGAGGCGCGGCTATGCGAGGCGATCTTCAGGAGAC 4597  
QY 4886 TGGAGCAGCACTACATGCGCGGATTAAGTCAAGCAATTAAGTCAACACAGAAATTTTACGC 4945  
Db 4596 TGGAGCGCGCGGCTATCCCGACCTGCGCAAGCACCTGACCGTCAAGCGCATCTTACGCC 4537  
QY 4946 CGTTTATTTTCGGAACAGCTGATGCCATCAGCGCTCGCGCTTTCGCTGGAGCGCA 5005  
Db 4536 CCGCGATTTTCAGCACCGAATGTCGCGGCGGCTATCGGCGAGCGCTTTCGCTGCGAGCGCA 4477  
QY 5006 TTTTGAACCAAGCGCGCTTTCGCGCGCATACCGGATGCGGATATCAGCAATCTCT 5065  
Db 4476 TCCTGACGCAATTCGCGCTTTCGCGCGCATTAACCGGCAACGCGGATCCGGAATCTCT 4417  
QY 5066 ATCTGCTGGTGGCGGTACGCAATCAGCGCGGCGGCTGCCCGGCTGATTCGCTTGGCGCA 5125  
Db 4416 ACATGCTGGGCGGCGACGATCCGCGTGGCGCATCCGCGTGTCTGCTGCGAGCGCA 4357  
QY 5126 AGGCACCGCGGCTGATGCTGGAGGATCGCGGATGATCGACAGCCTTTACTTGA 5185  
Db 4356 AGGCACCGCGGAGGTCTGCTGTCGAGCTTGGCGCTGCGATGACCGGATTCGACGCGGAC 4297  
QY 5186 GCAAGTAAACGCAAAACATGCGGCTGGGCTCGAAAGAGTTTTCGCCACCGCGCGCAAGCTGT 5245  
Db 4296 TTCCGAGGG---GCCATCGCGCAGGTTTCGCAAGCTTCGCGCAGGCGGCGCAAGCTGAT 4240  
QY 5246 TGATGACCGGCGCGGACGACGCTGATGCTGTATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5305  
Db 4239 GCGCGCGCGCATCCGCGAGGATACGCTCATGCTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4180

QY 5306 TGTGATTGATGGGCAAAACCTCGGCGAAGGCGGACGACGATGCGTCAAGACGCGCA 5365  
Db |||||  
QY 4179 CGTGATCGACGGGCGAGGTGATGGTTCTGCCCGGAG-----GGGGGCGGCGACCA 4126  
Db |||||  
QY 5366 GGCAGTATGACGATCTGCAAAATGAACCCGCGCGCTACAGCGGCGGCACATGGA 5425  
Db |||||  
QY 4125 GCGCGGCTGGGGGCGCTCGCGCCGACACGCTGGCGCGCTGACAGAGACGGCCGAT 4066  
Db |||||  
QY 5426 TGAACCGGCGTTTAGGGCGTTTCAGGAAGTGGCGATCATTCACAGCTGCGCAACAAC 5485  
Db |||||  
QY 4065 GTCGCGCGCTTCGCGGCGCTCGCCAGGTGCGCGGCGCATGATTTCCCGGACCTTTG 4006  
Db |||||  
QY 5486 GCGGTTTGATCTGGAAGGCTTCGCTATGATGACGCAACGATTCAGCGAGCTT 5545  
Db |||||  
QY 4005 GCCGATGGAACCTGATCGAGGGTTTCGCGATGATGTGCGGATCGCGATACCGAGCCT 3946  
Db |||||  
QY 5546 CGATGACACGCTGCGTTACTGCTATCACGTGCGGCGGTGCGTTCGATGATGCGCGG 5605  
Db |||||  
QY 3945 GGATGACGTGCTGGAATATCTTACACGTGCGGGGCTGCGGGGATGATGCGCGG 3886  
Db |||||  
QY 5606 CTAATAGGCGTGCAGCAAGAGCGTGTCTGATCACGCTCGCATTTAGGACTGCGGTT 5665  
Db |||||  
QY 3885 GGTGATGGCGTGCAGGACGATGCGGTGCTGATCGCGCTCGCATCTGGGCTTGGTT 3826  
Db |||||  
QY 5666 CCAGCTCACTAATTCGCGCGACATTTAGAAAGTCCGAAATGCTGCTGCTATCT 5725  
Db |||||  
QY 3825 CCAGCTGACAAATCGCTCGGACGATGATCGACGATGCGGCGCTGCTATCT 3766  
Db |||||  
QY 5726 GCGCAATCTGCTCGATCAGCGGGATTACGCGCGGATGCTGACTGCACCGCAACA 5785  
Db |||||  
QY 3765 GCCTGCGACTGGCTGCGCGAGCGGG-----GGCGAGCTTGGAGGTGCGTGC 3715  
Db |||||  
QY 5786 TGTGCGAGCGCTCGCTCACTGCGACGCGTTTAGTGGCGGAGCGGAACCTATTATCA 5845  
Db |||||  
QY 3714 TTGCGACGCGCTTATTCGTGATCATCGCTGCTTGACGCGCGCGCTTATATGC 3655  
Db |||||  
QY 5846 CTCGGCGCATCCGGTTTACCGGTTTACCGTGTGCTGCGGCTGGGCGCATCGTACGCG 5905  
Db |||||  
QY 3654 CTCGGCGCGGAGGGGCTTCCGATCTGCGCGCGCTGCGGCTGCTGATCGCGCGCG 3595  
Db |||||  
QY 5906 TCGCGCGTTTATCGGAAATTTGGCTCAAGTTGACGACGCGGCTGTGACGCGCTGGA 5965  
Db |||||  
QY 3594 GCTGCTATCTATCGCGCAATCGGACGCGCATTCGCGAGGTTGCGCGCGGCTATCG 3535  
Db |||||  
QY 5966 TTCGCGCAGCGCACAGTAAAGGTGAATAAATCGGCTGCTGCTGTAAGAGGGCGAGTTT 6025  
Db |||||  
QY 3534 CCAGCGGATCAGACGTCGAAGGCTGCCAAGATCGGGCTTCTGGCGCGGAGGCTTGA 3475  
Db |||||  
QY 6026 GCGGATCACTTCGCTGTGCTGCTGTAACCGCGTCCGGCTGCTGTGTGGCAGGCTCC 6085  
Db |||||  
QY 3474 CCGCGCGCATCGCGCTCGCGCGCGGCGGAATCAGCCGCGAGCTGTGTGGACCGGACC 3415  
Db |||||  
QY 6086 TCG 6088  
Db |||||  
QY 3414 GCG 3412

## RESULT 14

AAV40151/c

ID AAV40151 standard; DNA; 11233 BP.

XX

AC AAV40151;

XX

DT 10-AUG-1999 (first entry)

XX

DE DNA sequence of plasmid pZe4.

XX

KW Carotenoid; pigment; canthaxanthin; R1534; crtB; prephytoene synthase;

KW crtI; phytoene desaturase; crtY; lycopene cyclase; crtW2396; pZe4;

KW beta-carotene beta-oxygenase; food product; fermentation; ds.

XX

OS Synthetic.

XX JP10155497-A.  
XX 16-JUN-1998.  
XX 02-DEC-1997; 97JP-00348653.  
XX 02-DEC-1996; 96EP-00810839.  
XX (HOFF ) HOFFMANN LA ROCHE & CO AG F.  
XX WPI; 1998-391048/34.  
XX Preparation of carotenoid - comprises fermentation with transformed cell.  
XX Example 2; Fig 42-53; 80pp; Japanese.  
XX The invention describes the preparation of carotenoid pigments e.g. canthaxanthins using a cell transformed by a vector having DNA sequences (a) to (e) or substantially homologous sequences. (a) a DNA sequence (crtB) coding GPP synthase of Flavobacterium sp. R1534; (b) a DNA sequence (crtB) coding prephytoene synthase of Flavobacterium sp. R1534; (c) a DNA sequence (crtI) coding phytoene desaturase of Flavobacterium sp. R1534; (d) a DNA sequence (crtY) coding lycopene cyclase of Flavobacterium sp. R1534; and (e) a DNA sequence (crtW2396) coding beta-carotene beta-oxygenase of a microbe E-396 (FERM BP-4283). The carotenoid or a carotenoid mixture can also be used in preparation of food products. The method is an improved method of fermentation for carotenoid production  
XX Sequence 11233 BP; 2142 A; 3525 C; 3605 G; 1960 T; 0 U; 1 Other;  
SQ

Query Match 14.8%; Score 1037; DB 2; Length 11233;  
Best Local Similarity 56.6%; Pred. No. 1.4e-263;  
Matches 2038; Conservative 0; Mismatches 1535; Indels 30; Gaps 5;  
QY 2486 GTGCAGCAGCAAGTGTGCTGGCGAGCGGACCTGATGCGCAGCAGCATACGATGTGATT 2545  
Db 7666 GGGCGCGTGTGAAAGGACCGCGAAGGGGCGGATCGCAATACATGAGCCATGATCTGCTG 7607  
QY 2546 TTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2605  
Db 7606 ATCGCGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7547  
QY 2606 CAATGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2665  
Db 7546 GATGCGCGATCGTATGCTGCGCGCGTCCGCGCCCTCGGACGACGACACCTGCTGCTGCTG 7487  
QY 2666 TTTTCATCAGCGATCTCAGCGCGCAACAACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2725  
Db 7486 TGCCACGACAGGATCTTTCGCGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7427  
QY 2726 CGTTCGTCAGGTTATCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2785  
Db 7426 GATGCGCGATCAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7367  
QY 2786 TGTTCATCGCATCAGCGGATTTTTCGCGCGGATTTTTCGCGCGGATTTTTCGCGCGGATTTTTCG 2845  
Db 7366 GGTTCGATCGAGGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7316  
QY 2846 TGGCAAAACACAGCGCTACACAGGTAACCCAGCAGGTCAGCTGCGGATGCGGCTGCGGCTGCG 2905  
Db 7315 CGGTGGAATACGATGTGCTGCGACGCTGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 7256  
QY 2906 GAACTTGTGCGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2965  
Db 7255 CGGATCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7196  
QY 2966 CTGGGTTATCAGGTTTCTTTCGCAAGAGTGGAGCTGGCGAGCCGCGCGGCTGCGAG 3025  
Db 7195 GTGGGTTTCCAGAAATTCGTGGCGCTGAGATCGAGACCGACGCGCGCGCGCGCGCGCGCGCG 7136  
QY 3026 CAGCCGATCTGATGATGCCACCGCTGATCAGCAAGCGGGTTATGCTTTTCTTCTACAGC 3085

Db	7135	CGCCGATGATCATGACGACGACGCTTGGCAGATGAGCGGTACCGCTTCACTATCTG	7076	6058	CCACGCCCTTCTGAAGCTGGCCACAGATGTGAACGCCGCCGCCGCTGATGCGCCTGC	5999
Qy	3086	CTGCCGCTCAGCGCGATCGCTATTGATTAAGATACCAATTAAGTTAAACGACCCCGCG	3145	4166	AAGATGCGCGAGCGTCTACAGCATGGTGGGAAATTTATTCAGGACGATCATCTGCGTC	4225
Db	7075	CTGCCCTTCACTCCACCGCATCTGATCGAGGATACGCCGTACAGCGCGCGCGAT	7016	5998	AGGCATACCGCTCGGTCCACAGCATGGTGGCGCGCTTATCCAGGACCCGATCTGCGGC	5939
Qy	3146	CTGGCGGAGAACACCGCTCTGAGCACATCGCCGACTATGCAATACAGCAAGGCTGAGCG	3205	4226	AGSCGTTTTCTTCCACTCATTTGCTGGTGGCGGTAATCTTTTGCATCAAGTCAATCAT	4285
Db	7015	CTGAGAGATGGCGCTGGCGAGCGGCTGCTGGAATATGCGCGAGCGGGCTGGACC	6956	5938	AGSCCTTCTGTTCCACACGCTGCTGGTGGCGGGAACCCGTTTTTCGACCACTCATCT	5879
Qy	3206	CTGAGTACGCTGCTGTTGAAGACAGCACGGCATATTACGATTAACCTGAGCGGCATC	3265	4286	ATACCTTAATCATGCGCTGGAGCGTGAATGGGCGTGTGGTTTTCCGCGCGCGACCG	4345
Db	6955	GGGAGAGATGCGCGCGAAAGGGGCATCTGCCCATCGCGCTGGCCCATGACGCCATA	6896	5878	ATGCGCTGATCCATGCGCTGGAACCGCGCGCGCGGCTGCTGGTCGCAAGGCGGCACA	5819
Qy	3266	GATCGATTCTGGCMAACAGCAGCGCGCAAGGTGCGAGCGGCTGCGCGCGGGCTGTTT	3325	4346	GCGCGTGGTGCAGGCGCATGCTTTCGAGGACTTGGGCGCGGAGCTGTTACTGA	4405
Db	6895	GGCTTCTGGCGGACACCGCAGGGGGCGGTGCGGTTGGCTGGGGCAGGGCTGTTT	6836	5918	ACAGCTGGTGGCGGATGCTCGCCCTGTTTCGAGCGTCTTGGCGGACGCTGCTGA	5759
Qy	3326	CATGCCACCAACCGGTTACTCTTGGCGTCCGCGTGGGCTAGCGGATGTTGTCAGCG	3385	4406	ATGCCGAAGTGAAGCCAGCTGGAAACACAGCGGCAATCGCATTTAGCGCGTTCAGTTAGAG	4465
Db	6835	CACCCGTCACCGGATATTGCTGCGCTATGCGCGGAGTGGCGATGCGGATGCCATCGCGG	6776	5758	ATGCCCGCTCACCGGATCGACACCGAGGGGATCGCGCACGGGCGTCAOGCTCTGG	5699
Qy	3386	CTGTTGCCACCGATGCCCTCAGCTCAGCCACATATCGAAGCTTTGCCGCTCAGCAG	3445	4466	GCGGACGACCTTCGATGTCGCGCGCTTCCAAATGCGGACGTGGTGCATACCTACG	4525
Db	6775	CGCGACTGACGACCGCTCGCCCGCTGCGCGGCTGCGCGGCTGGCCCATCATCGCGG	6716	5698	ACGGCGGCGATTGCGCGCGGATACGTTGGCCAGCAACGGCGACGTGTCACAGTATC	5639
Qy	3446	TGCGGGAACAGCGATTTTTTCCTCTGCTAAACCGCATGCTGTTTTTGGCGGTAAGCG	3505	4526	ACAACTGCTTCGCCACCATCCCTGCGCAATGAACGTGCGCATCGCTGAAGCTAAGC	4585
Db	6715	GATCGCG--ACCGCTTCTGGGCTGCTGMAACGGATGCTGTCGCGGCTGCCCGGCC	6659	5638	GCGACTGCTGGGCCATACCGCGCGGGCGCACCAAGGCGCGATCTCGAACCGCAGC	5579
Qy	3506	CAGAGCGCTGGCGGCTGATGAACGTTTTTACCGGCTCGATGCGGGTTAATAGCGC	3565	4586	GCATGAGCAACTGCTGTTTGTACTCTATTTTGGCTGATATCAGCGCATGAACAGCTCG	4645
Db	6658	GACCGTGCATCGCCCTGTCAGCGGTTCTACCGCTGCGCGAGCCGCTGATCGAGCG	6599	5578	GCTGGTGCATGCTGCTGTTTGGCGCGGTTATGCTGATTTGATGATGAGATTTTCAACA	4705
Qy	3566	TTTTTACCGCGGCAACTGCGCTCGCGATMAAACCGGATTTCTGTGGCGAAGCGCG	3625	4646	CGCACCAACACGCTGTTTTTGGCGCGGTTTACGAGGCGCTGGTGAACAGATCTTCAACG	5459
Db	6598	TTCTATGCCGGGCGCTGACATTTGGCGGCGGCTGCGATCGTCAACCGCGCGCGCC	6539	4706	GCAGCAGCTGGCAGAGATTTTTCATTTTACCTGCGAGCGGCTGCGAGCGGATCCGT	4765
Qy	3626	GTGCCCATCGGTGAAGCGCTGGCGGCTGTTGAATTTCTGTAACAGGAGAGAGCA	6479	5458	GGCCACGCTCGCGGAGATTTTCTCGATGATATCTGCATTCGCTCGCTGCGCATCCCA	5399
Db	6538	ATTCGCTGTGCGAGCGCTGCTGCTGCCGAAACGCGCCCTGCTGCGAGGAGAGCA	6479	4766	CGCTGGCACCGCCCGCTGGCGAGCTTTATGTTGTTAGCGCGGCTGCGCATCTCGCA	4825
Qy	3686	TGAACCGCATTTGATTTGGCGAGCGCTTTGGCGGCTGGCGCTGGCGATTCGCTGC	3745	5398	GCCTGGCCCCCGAGGGGATGTCAGCATTAAGTCTTGGCGCGCTTCCGATCTGGGCC	5339
Db	6478	TGATTCGCCCATCGTCAATCGCGCGAGTTTCGGCGGCTTGGCTTGCATCCGCTGC	6419	4826	CCGCTGCAATCACTGGCAACAGGAAGCAACCGCGCTTGGCGGATCGAATTTTCTTATC	4885
Qy	3746	AAGCGGGGCGATACCAACCACTTACTCGAGCAGCGGCAAAACCGGGGAGCGGCT	3805	5338	GCGCGATGTCGATTTGGGAAGCGAGGCGCGGCTATGCGCGAGCGCATCTTCAGGAAC	5279
Db	6418	AATCGCGGCGATCGGACCAACCATCTGTCGAGGCGCGCGACAAAGCGCGCGCGCT	6359	4886	TGGACGACACTACATGCGCGGATTAAGTCAAGCAATTAAGTACACAGATGTTTACGC	4945
Qy	3806	ATGTTGTTGAGCAGTGGCTTTTACCTGATGCGGACCGACCGGCTGATCACCGATCCCA	3865	5278	TGGAGCGCGCGCCATCCCGGACCTTGGCGAAGCACTGACCGTCAAGCGCATCTTCAGCC	5219
Db	6358	ATGTTGGAACGATCAGGCGCACTCTTCGATGAGCGGCGCGAGCTGTCGACGCCCG	6299	4946	CGTTTGAATTTTTCGCGACACGCTGCAATGCCCATACGCTCGGCTTGGCGTGGAGCGCA	5005
Qy	3866	GCGCATCAAGAGTTGTTACGCTGGCAGGAAATCGCTCAGCGATTAAGTCAAGCTGA	3925	5218	CGCGCATTTTCAGCACCGCACTGTCGCGCCATACCGGAGCGGCTTCTCGGTTCAGGCGCA	5159
Db	6298	ACAGCTCGAGAGCTGTGGGCGCTCAGCGGCGCAACGATGGAGCGTGAAGTGAAGCTG	6239	5006	TTTTTGAAGCAAGCGGCTGTTTCCCGCGCATTAACCGGATGCGGATATCAGCAATCTCT	5065
Qy	3926	TGCCGGTAAAGCCCTTCTATCGCTGCTGGAGAGTGGCAACAGCTTGAATACGACA	3985	5158	TCTGACGCAATCCGCTGTTCCCGCGCATAAACCGCAACCGCGCATCCGAACTTCT	5099
Db	6238	TGCCGCTTCGCGCTTCTACCGGCTGACATGGCGGAGCGCGCGAGCTTGAATACGTGA	6179	5066	ATCTGTTGGTCCGCTGATCGCATCCAGGCGCGGCGTCCCGGCGTGAATCGGTTCCGCA	5125
Qy	3986	ATAATCAGCGCTGCTGAGCAGAGATCGCACTTCAATCCGCAAGATGTAAGGCT	4045	5098	ACATCTGGGGGGGCGGCGCATCCGGGTGGGCGATCCCGGCTGCTGTCGACGCGCA	5039
Db	6178	ACAGCAGCAGCAGCTGATCGCCAGCTGCTCTCTTCAATCCCGCGGATGTCGATGCT	6119	5126	AGGCCACCGCAGGCTGATGCTGGAGGATCGCGCGCATGATCCGACAGCTTACTTGA	5185
Qy	4046	ATCGTCAATTTCTGCTTATTCAGCTGAAGTATTATAGAGGGTTATCTGAACTCGCA	4105	5038	AGGCCACCGCGAGGCTATGCTGTCGACCTGCGCGTGCATGACCGATCTCGCGCGAC	4979
Db	6118	ATCGCGCTTCCACGATTAACCGGAGGAGGCTATATCGGAGGGGTTATCTGAAGCTGGGA	6059	5186	GCAAGTAAACGCAACCATGCGGCTGGGCTCGAAGAGTTTTCGCCACCGCGCGCAAGCTGTT	5245
Qy	4106	CGGTGCGGTTCTGAGAGTGGCTGATCATGCTGCGCGTCCGCGCGAGTTGGGAGCTGTC	4165	4978	TTCCGAGAGCG--GCCATCGCGAGGTTTCGCAAGCTTTCGCGGAGGCGCGCAAGCTGAT	4922



DR P-PSDB; AAW06513, AAW06514, AAW06515, AAW06516, AAW06517, AAW06518, AAW06519.

XX Flavobacterium gene sequences encoding carotenoid biosynthesis enzymes -

PT For the production of carotenoid(s), useful in foods and animal feeds.

XX Example 2; Fig 7; 80pp; English.

XX Genomic DNA (AAT45143) of Flavobacterium sp. R1534 includes genes of the carotenoid biosynthesis pathway. The sequence was deduced from inserts of 6 clones obtained from genomic libraries e.g. by PCR amplification (see also CC AAT45144-45) and use of partial clones to screen the library. The CC identities of the gene products (see also AAW06513-19 and AAW06517) were CC ded. by examining carotenoid accumulation in E. coli hosts transformed CC with deleted variants of the gene cluster. The isolated genes can be used CC in different combinations to produce carotenoids in transformed host CC cells. (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 8625 BP; 1458 A; 2898 C; 2964 G; 1295 T; 0 U; 10 Other;

Query Match 14.8%; Score 1035.4; DB 2; Length 8625;

Best Local Similarity 56.5%; Pred. No. 3.3e-263;

Matches 2037; Conservative 0; Mismatches 1536; Indels 30; Gaps 5;

QY 2486 GTCCAGCAAGTCGTGCTGGCGGAGGACCTGATCGCGACCAATACGATGATT 2545

DB 6984 GGGCGGCTGCTGAAGGACCGGAGGGGCGGATCGCAATATAGCCATGATCTGCTG 6925

QY 2546 TTGCTGCTGCTGCACTGGGAATGGCTTGAATGCGCTGCTGCTGCTCAATTCGAGCCA 2605

DB 6924 ATCGGGGCGGGGCTGCTCGGTGCTGATCGGCTTGGCGCTTGGCGACCGAGACCG 6865

QY 2606 CAACCTGAATGCTGTTGCTGGAGAGCGATGCGCATCCGGCAGGCAATACCTGCTGCG 2665

DB 6864 GATCGCGCATCGTATGCTCGACGGCGGCTCGGCCCTCGGACAGCACACCTGGTCC 6805

QY 2666 TTTTCATCAGCGCATCTCAGCGCGCAACTTTCGCTGGCTGCAACCGCTGATTAACCGTG 2725

DB 6804 TGCCACGACAGGATCTTTCGCGGAATGCTGGCGGCTGTCGCCATTCGTCGCGC 6745

QY 2726 CGTTGGTCAGGTATCAGGTGCGTTTTCCTGCGCTGCGCGCAATCTGACGGGGATTAT 2785

DB 6744 GAATGGACGATCAGGAGTTCGCTTCCGACCAATTCGCGCGCTGACGACAGCTAT 6685

QY 2786 TGTTCATCGCATCAGCGATTTTTCGCCCATCTTACGCGCGATGGGTGACGATCTG 2845

DB 6684 GGCTCGATCAGGCGGCGCGCTGATCGGGTGTGCA-----GGGTGCTGATCTG 6634

QY 2846 TGGACAAACAGCGTACACAGGTAAACCCACGACGATGACGCTGGCGGATGGCGGT 2905

DB 6633 CGGTGGAATACGATGTCGCGACGCTGGAAGATACCGCGCGACGCTGACGGGACGCTCG 6574

QY 2906 GAACCTGTGCGGAAGTGTGATGATGCTGGCGCTGACGCGCGCCACATCTGCGAG 2965

DB 6573 CGGATCGAGGTGCTGCTGATCGACGCGGCTGGCGCGTGGTCGAGACCCCGACCTGACC 6514

QY 2966 CTGGGTATCAGGTGTTTCTTGGACAAGAGTGGCAGCTGGCGGACGCGCGCTGCGAG 3025

DB 6513 GTGGGTTTCCAGAAATTCGTGGCGCTCGAGATCGAGCCGACGCGCCCGATGGCGTGG 6454

QY 3026 CAGCCGATCTGATGGATGCCACCGTTCGATCAGCAATGCGCTATGCGGATGCTGACAG 3085

DB 6453 CGCCCGATGATCATGGACGCGACCGTTCCGAGATGGAAGGATACCGGTTCATCTATCTG 6394

QY 3086 CTGCGCTCAGCGCGGATGGCTTGAATGAAGATACCATTAACCGACCGCGCG 3145

DB 6393 CTGCGCTTCACTGTCACCGCATCTGATCGAGGATACGCTACAGGACGCGGCGGAT 6334

QY 3146 CTGGGCGGAAACACCGCTCGTCAGCAATGCGCTATGCGATGCGAATACGAGGCTGAGCG 3205

DB 6333 CTGGACGATGGCGCTGCGCGAGGCTCGCTGGACTATGCGCGAGCGGGGCTGGACC 6274

QY 3206 CTGAGTACCTGCTGGTGAAGACGACGCGCATTAACCGATTACCTGAGCGGCAATC 3265

DB 6273 GGCAGGATCGGCGGGAAGGGCATCTGCCCATCGCGTGGCCCATGACCCATA 6214

QY 3266 GATCGAATTCGCAACAGCAGCGCGCCAGCGTGCAGCGGCTCGCGCGGGCTGTTT 3325

DB 6213 GGTTCGCGGACCAACGCGGAGGGGCGGTGCGCGTGGCTGGGGCAGGCTGTTTC 6154

QY 3326 CATGCCACACCGGTTCCTTTCGCGTCCGCGCTGGCGCTAGCGGAGTTGTAGACGG 3385

DB 6153 CACCCGTACCGGATTTGCTGCTATGCGCGCAGGTCCGATCCCATCGCGCG 6094

QY 3386 CTGTTGCCACCGATCCCTCAGCTCAGCAACATATCGAAGCTTTGCCGCTCAGCAG 3445

DB 6093 CGGACTGACGACCGCTCCGCCCTCGCGGCTGGCGCTGCGGCAATCGATCGCGC 6034

QY 3446 TGGCGGACAGGATTTTTCGCTGCTAAACCGCATCTGTTTGGCGCGTAAAGCG 3505

DB 6033 GATCGCG---ACCGCTTCTTCGCGCTGCTGACCGGATCTGTTCCGCGCTGCCGCC 5977

QY 3506 CAGCAGCGCTGGCGCTGATGCAACGTTTTCACCGCTCGATGCCGGTTAATAGCCG 3565

DB 5976 GACGCTGCTATCGCTGCTGACGCGTTCTACCGCTGCGCGCAGCTGATCGAGCG 5917

QY 3566 TTTTACCGCGGCACTGCGCTCGCGGATAAACCGGATTTCTGTGCGCAAGCGCGCG 3625

DB 5916 TTCTATGCGGGCGCTGACATTTGGCGACCGGCTTCGATCTGATCGGACGCGCGCC 5857

QY 3626 GTGCCCATCGTGAAGCGCTGCGCGCTGTTGAATTTCTGTGAAACGAGGGAAGAAAA 3685

DB 5856 ATTCCGCTGTCGAGCGCTGCGCTGCTGCCCGAACGCGCTGCTGACGAGGAGCA 5797

QY 3686 TGAACCACTATGATTTGGCGCAGCTTTGGCGCTGCGCTGCGGATTCGCTGCG 3745

DB 5796 TGAGTTCCGCTATCGTATCGCGCAGGTTTCGCGGCTTGGCTTGCATCGCCTGC 5737

QY 3746 AAGCGGCGGATACCAACCACTTACTCGAGCAGCGGACAAACCGGGCGGACGCGCT 3805

DB 5736 AATCGCGCGGATTCGCGACCACTCTCGAGGCGCGGACAAAGCCGCGCGCGCT 5677

QY 3806 ATGTGTTTGAAGCAGTGGCTTTTACCTTCGATGCGGACCAACGCTGATCAGCATCCCA 3865

DB 5676 ATGTCTGGAACGATCAGGCGCACGCTTTCGATGACGCGCGGACGCTGACCGCCCG 5617

QY 3866 GCGCATCGAAGTGTTCACGCTGCGGAGGAAATCGCTCAGCGATTTAGCTCAGCTGA 3925

DB 5616 ACAGCTCGAGAGCTGTGGGCGCTCAGCGGCGCAACCGATGGAGCGTGACGCTGC 5557

QY 3926 TGCCGCTAACGCTTCTATCGCTGCTGGGAAGATGGCAACAGCTTGATTAACGACA 3985

DB 5556 TGCCGCTCTCGGCTTCTACCGCTGACATGGGCGGACGCGCGGATTCGAATACGTA 5497

QY 3986 ATAAATCAGCGCTGCTGAGCAGCAGATCGCACGCTCAATCCGCAAGATCTAGAAGCT 4045

DB 5496 ACAGCAGCAGCAGCTGATCCGCGCAGGTGCGCTTCAATCCGCGGATTCGATGGCT 5437

QY 4046 ATCTGCAATTTCTGCTTATTCACGCTGAAGTATTTAGAGAGGTTATCTGAAACTCGGCA 4105

DB 5436 ATCGCGCTTCAACGATTCGCGGAGGTTCTATCGAGGAGTTCTGAAGCTGGGA 5377

QY 4106 CGGTGCGGTTCTGAGGTGCTGATGCTGCGGCTCGCGCGCAGTTGGAGCTGCTGC 4165

DB 5376 CCACGCGCTTCTGAAAGTGGCGCAGATGCTGAACCGCGCGCGGCTGATGCGCCTGC 5317

QY 4166 AAGCATGGCGGCTGATACAGCATGGTGGCGAAATTTATTCAGGACGATCATCTGCGTC 4225

DB 5316 AGGCATACCGCTGCTGCGGCTCAGCATGGTGGCGGCTTCAATCAGGACCGCATCTGCGGC 5257

QY 4226 AGCGGTTTCTTCCATCATTTCTGTTGGGCGGTAATCTTTTGAAGCTCATCTGACT 4285

DB 5256 AGGCTTCTGTTCCACAAGCTGCTGGTGGCGGGAACCGGTTTTCGACAGCTGATCT 5197

QY 4286 ATACTTAAATTCATGCGCTGAGCGTGAATGCGGCTGTTGTTTCCGCGCGCGGACCG 4345



Db 5196 ATGGCGTATCATGGCTGGAAACGGCGCGGGCGTCTGGTTCCGCAAGGGCGGACCA 5137  
Qy 4346 GCGCGCTGTGACGGCATGGCGCGCATGTTTCAGAGACTTGGCGCGCGAGCTGTTACTGA 4405  
Db 5136 ACCAGCTGTGCGGGCATGTGCGCCCTGTTCCAGCGTCTTGGCGGACGCTGCTGTA 5077  
Qy 4406 ATGCCGAGTGAAGCAGCTGGAAACAGCGGCAATCGCATTTAGCGGCTTCAATTAGAG 4465  
Db 5076 ATGCCGCGTACGGGGATCGACACCGAGGGCGATCGGCCACGGGCGTCACTGCTGG 5017  
Qy 4466 GGGGACGACGCTTCGATGCGCGGCTGTGGCTTCAATGCCAGCTGGTGATCACTACG 4525  
Db 5016 ACGGGCGGAGTGGCGCGGATACGGTGGCCAGCAACGGCGACGTGATGACAGTATC 4957  
Qy 4526 ACAACCTGTTCGCCACCAATCGCTGGCAATGAACGTGCGCATCGCTGAAGCGTAAGC 4585  
Db 4956 GCGACCTGTGGGCCATACCCCGCGGCGCACCAAGCGCGCATCTCTGAACCGGCGAC 4897  
Qy 4586 GCATGAGCAACTCGCTGTTGTACTTATTTTGGCTGATCAGCGGATGACAGCTCG 4645  
Db 4896 GCTGTCGATGTCGCTGTTGCTGCTGCAATTCGGCTGTCTCAAGCGCCCGGAGAACCTGG 4837  
Qy 4646 CCGCACACCTGTTGTTGGCCCGGCTTATCGTGAGTTGATCGATGAGATTTTCAACA 4705  
Db 4836 CCCACACAGCTCATCTTCGGCCCGCGCTACAGGGCTGTGAAAGATCTTCAACG 4777  
Qy 4706 GAGCGAGCTGCGACAGATTTTCACTTTACCTGCACCGCCCTCCAGCAGCATCCGT 4765  
Db 4776 GCGCACGCTGCGGAGATTTCTCGATGATCTGCAATTCGCCCTGGGTGACCGATCCA 4717  
Qy 4766 CGCTGGCACCGCGCTGCGGAGCTTTATGTTAGCGCGGTGCGGATCTCGGA 4825  
Db 4716 GCTTGGCCCCCGAGGGGATGTCACCGCAATACGTCCTTGGCGCGTTCGCGATCTGGCC 4657  
Qy 4826 CGCTGACATCGACTGGCAACAGGAAGACCGCTTGGCGATCGAATTTTGTCTATC 4885  
Db 4656 GCGCGATGTCGATTTGGGAGCGGAGGCCCGGCTATGCCAGCGCATCTTCGAGGAC 4597  
Qy 4886 TGGAGCAGCATACATGCGGGATTAAGTACAGCAATTAAGTACACACAGAAATTTTACGC 4945  
Db 4596 TGGAGCGCGCCATCCCGACCTGCGCAAGCACCTGACCGTACAGCGCATCTTCAGCC 4537  
Qy 4946 CGTTTGAATTTCCGACACGCTGCATGCCCATCACGGCTCGGCTTTCGCTGGAGCGCA 5005  
Db 4536 CGCGCGAATTTACGACACCGAATGTCGCGCCCATCACGGCAGCGCTTCTCGGTGAGCGCA 4477  
Qy 5006 TTTTGAACGAAAGCGCTGTTCCGCGCGCATAAACCGGATGCCGATACGCAATCTCT 5065  
Db 4476 TCTGACGCAATCCGCTGTTCCGCGCGCATAAACCGGATGCCGATCCCGAATCTCT 4417  
Qy 5066 ATCTGTTGGTCCGCTACGATCCAGGCGCGGGGTGTCGCGGTGATTCGGTTCCGCCA 5125  
Db 4416 ACATCGTGGGGCGGCGACGATCCGGGTGCGGGCATCCCGGTGTGTTGGACGGCA 4357  
Qy 5126 AGGCGACCGCAGGCTGATGCTGGAGGATCGCGCGGATGAATGACAGCCCTTACTTGA 5185  
Db 4356 AGGCGACCGCAGGCTGATGCTGTGCGACCTGGCGCTGCATGACCGCATCTGACGCGGAC 4297  
Qy 5186 GCAAGTAACGCAAAACCATGGCGGTGGCTCGAAGAGTTTCGCCACCGCGCGCAAGCTGT 5245  
Db 4296 TTCCGAGGCG---GCCATCGCGAGGTTTCGCAAGCTTTCGCGAGCGCGGCAAGCTGAT 4240  
Qy 5246 TGATGACCGACGCGCGCAGCATGCTGATGCTGTATGCTGTGCTGCTGCTGCTGCTGCTG 5305  
Db 4239 GCGCGCGGCGCATCCGCGAGGATACGGTCATGCTCTATGCTGTGTGAGGATGCGGATGA 4180  
Qy 5306 TGTGATTGAGGCAAAACCTGGCGGAGAGCGGCGACGAGCATGCGTGCAGAGACGCGCA 5365  
Db 4179 CGTGATCGACGGGCGAGTGTGGGTTCTGCCCGCGAG-----GCGGCGGCGGACCCACA 4126  
Qy 5366 GGCAGTATGAGCATCTGCAATTTGAACCGCGCGCTACAGCGCGCGCATGGA 5425  
Db 4125 GCGCGGCTGGGGGCGCTCGCGCGCGACACGCTGGCGCGCTGTCAGAGGACGGCCCGAT 4066

Qy 5426 TGAAACCGCGTTTAGGGCTTTACGAAAGTGGCGATCATTCACAGCTGCGCGCAACAAT 5485  
Db 4065 GTCCGCGCTTTCGGCGGCTGCGCAGGTCCCGGGGCGATGATTTCCCGGACCTTTG 4006  
Qy 5486 GCGGTTTTCATCATCTGGAAGGCTTCGCTATGAGATCAGCAGCAACATTTAGCGAGCTT 5545  
Db 4005 GCGGATGACCTGATCGAGGGTTTCGCGATGATGTCGCGGATTCGCGAATACCGAGCTT 3946  
Qy 5546 CGATGACACGCTGCTGCTTACTGCTATCAGCTCCGCGCGTGTGCTGCTTGTGATGATGCGCG 5605  
Db 3945 GGATGACGCTGCTGGAATATTCCTACACGTCCGCGGGTTCGTGGGCGTATGATGCGCG 3886  
Qy 5606 GGTAAATGGGCTGCGCGACGAAAGCGGTGCTGATCAGCCTCGGATTTAGGATGCGGCTT 5665  
Db 3885 GGTGATGGGCTGCGAGCAGATGCGGTGCTGATCGCGCTCGGATCTGGGCTTTCGCTT 3826  
Qy 5666 CCAGCTCACTAACATTTGCGCGCATTTGTAGAAGATCCGAAATGTCGCTGCTATCT 5725  
Db 3825 CCAGCTGACAAACATCGCTCGGACGTATCGCCCATCGGGCGCTGCTATCT 3766  
Qy 5726 GCGCAATCTCGCTCGATCAGGCGGGATTTACGCGCGGATACGCTGACCTGCACCGCAACA 5785  
Db 3765 GCCTGCCGACTGGCTGGCGGAGCGGG-----GGCGACGCTTGAGGTCGGTGCC 3715  
Qy 5786 TCGTGCAGCGCTCGCTCACTGGCAGCGGTTTGTGGCGGAGGCGGAACCTTATATCA 5845  
Db 3714 TTCCGACGCGCTTATTCCTCATCATCCGCTGCTTGACGCGCGCGAGCCCTATTTATGC 3655  
Qy 5846 CTCGCGCGATCCGCTTACCGGTTTACCGGTTTACCGCTGCGCTCGGCGTGGCCATCGTACGGC 5905  
Db 3654 CTCGCGCGGCGACGGGCTTTCGCGCTCTGCGCGCGCTGCGGCTGCTGATCGCGCGCG 3595  
Qy 5906 TCGCGCGCTTTATTCGCAAAATTTGGCGTCAAAAGTTTCAGCACGCGGCTGTGACGCTGGGA 5965  
Db 3594 GCTGCTATCTATCGCGCAATCGGACCGGCAATCGGAGGTTGGCCCGAGGCTATCG 3535  
Qy 5966 TTCAGCGCAGCGCACAGTAAAGGTGAAAACTGGCGCTGCTGCTGTTGAAAGGGGCGAGTTT 6025  
Db 3534 CCAGCGGATCAGACGCTGCAAGGCTGCCAAGATCGGGCTTCTGGCGCGCGAGGCTTGA 3475  
Qy 6026 GCGGATCACTTCGCGGTGTCTCGTCTGAAACCGGCTCGGCTGCTGTTGCGAGCGTCC 6085  
Db 3474 CCGCGCGGATCGCGCTTGGCGCGGCGGAAATCAGCCGCGACGGCTGTGAGCCGACCC 3415  
Qy 6086 TCG 6088  
Db 3414 GCG 3412

Search completed: January 25, 2005, 05:06:30  
Job time : 2953 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 25, 2005, 03:38:11 ; Search time 501 Seconds  
(without alignments)  
9929.763 Million cell updates/sec

Title: US-10-808-979-18  
Perfect score: 6999  
Sequence: 1 gcgatgcaatggaaatgt.....gacgttttagacatgcaata 6999

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*  
1: /cgn2\_6/prodata/1/ina/5A-COMB.seq:\*  
2: /cgn2\_6/prodata/1/ina/5B-COMB.seq:\*  
3: /cgn2\_6/prodata/1/ina/6A-COMB.seq:\*  
4: /cgn2\_6/prodata/1/ina/6B-COMB.seq:\*  
5: /cgn2\_6/prodata/1/ina/PCUTS-COMB.seq:\*  
6: /cgn2\_6/prodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2624.6	37.5	6918	1	US-07-783-705A-13
2	1035.4	14.8	8625	3	US-08-980-832-1
3	1035.4	14.8	8625	4	US-09-920-923B-1
4	1035.4	14.8	11233	3	US-08-980-832-27
5	1035.4	14.8	11233	4	US-09-920-923B-27
6	826.2	11.8	1479	1	US-07-783-705A-10
7	713.4	10.2	1518	1	US-08-095-726-7
8	713.4	10.2	1518	1	US-08-096-043-7
9	713.4	10.2	1518	1	US-08-096-623A-7
10	708.4	10.1	1522	1	US-08-095-726-9
11	708.4	10.1	1522	1	US-08-096-043-9
12	708.4	10.1	1522	1	US-08-096-623A-9
13	619.4	8.8	1482	3	US-08-660-645A-6
14	619.4	8.8	1482	3	US-09-298-718-6
15	619.4	8.8	1482	3	US-09-546-969-6
16	619.4	8.8	1482	4	US-09-547-267-6
17	515.6	7.4	1198	1	US-08-095-726-5
18	515.6	7.4	1198	1	US-08-096-043-5
19	515.6	7.4	1198	1	US-08-093-577-5
20	515.6	7.4	1198	1	US-08-096-623A-5
21	449.6	6.4	1083	1	US-08-331-004A-1
22	449.6	6.4	1083	5	PCT-US95-13937A-1
23	438	6.3	1235	1	US-08-095-726-13
24	438	6.3	1235	1	US-08-096-623A-13
25	430	6.1	1232	4	US-08-908-758-1
26	430	6.1	1232	4	US-09-023-587A-1
27	426.8	6.1	1235	1	US-08-095-726-15

28	426.8	6.1	1235	1	US-08-096-623A-15
29	419.8	6.0	1296	1	US-07-783-705A-8
30	414	5.9	891	1	US-07-783-705A-11
31	399.6	5.7	947	1	US-08-096-623A-16
32	399.6	5.7	947	1	US-08-096-623A-17
33	386.2	5.5	1149	1	US-07-783-705A-9
34	357.2	5.1	2886	1	US-08-663-310-7
35	357.2	5.1	2886	2	US-09-006-491-7
36	357.2	5.1	2886	3	US-09-335-919-7
37	346.2	4.9	909	1	US-07-783-705A-7
38	308.2	4.4	1157	1	US-08-095-726-3
39	308.2	4.4	1157	1	US-08-095-726-3
40	308.2	4.4	1157	1	US-08-096-043-1
41	308.2	4.4	1157	1	US-08-096-043-3
42	308.2	4.4	1157	1	US-08-093-577-1
43	308.2	4.4	1157	1	US-08-093-577-3
44	308.2	4.4	1157	1	US-08-096-623A-1
45	308.2	4.4	1157	1	US-08-096-623A-3

ALIGNMENTS

RESULT 1  
US-07-783-705A-13  
; Sequence 13, Application US/07783705A  
; Patent No. 5429939  
; GENERAL INFORMATION:  
; APPLICANT: Misawa, No. 5429939ihiko  
; APPLICANT: Kobayashi, Kazuo  
; APPLICANT: Nakamura, Katsumi  
; APPLICANT: Yamano, Shigeyuki  
; TITLE OF INVENTION: DNA SEQUENCES USEFUL FOR THE  
; SYNTHESIS OF CAROTENOIDS  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Ladas & Parry  
; STREET: 26 West 61 Street  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10023  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 720Kb storage  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: N/A  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/783,705A  
; FILING DATE: 19911023  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 1-103078  
; FILING DATE: 21-APR-1989  
; APPLICATION NUMBER: JP 2-53225  
; FILING DATE: 05-MAR-1990  
; APPLICATION NUMBER: US 07/519,011  
; FILING DATE: 19-APR-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Schwadron, Janet I.  
; REGISTRATION NUMBER: 33,778  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-708-1935  
; TELEFAX: 212-246-5959  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6918 base pairs  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: other nucleic acid (plasmid DNA)  
; ORIGINAL SOURCE: Erwinia uredovora

Wed Jan 26 08:20:42 2005

```

STRAIN: 20D3
CELL TYPE: unicellular organism
FEATURE:
NAME/KEY: CDS
NAME/KEY: /gene="crt E"
LOCATION: 225 to 1133
IDENTIFICATION METHOD: by experiment
OTHER INFORMATION: codes for geranylgeranyl
OTHER INFORMATION: pyrophosphate synthase
NAME/KEY: CDS
NAME/KEY: /gene="crt X"
LOCATION: 1143 to 2438
IDENTIFICATION METHOD: by experiment
OTHER INFORMATION: codes for zeaxanthin glycosylase
NAME/KEY: CDS
NAME/KEY: /gene="crt Y"
LOCATION: 2422 to 3570
IDENTIFICATION METHOD: by experiment
OTHER INFORMATION: codes for lycopene cyclase
NAME/KEY: CDS
NAME/KEY: /gene="crt I"
LOCATION: 3582 to 5060
IDENTIFICATION METHOD: by experiment
OTHER INFORMATION: codes for phytoene desaturase
NAME/KEY: CDS
NAME/KEY: /gene="crt B"
LOCATION: 5096 to 5986
IDENTIFICATION METHOD: by experiment
OTHER INFORMATION: codes for phytoene synthase
NAME/KEY: CDS
NAME/KEY: /gene="crt Z"
NAME/KEY: /note="on the complementary strand"
LOCATION: 5925 to 6452
IDENTIFICATION METHOD: by experiment
OTHER INFORMATION: codes for @-carotene hydroxylase
PUBLICATION INFORMATION:
AUTHORS: Misawa, No. 5429939ihiko
AUTHORS: Nakagawa, Maaya
AUTHORS: Kobayashi, Kazuo
AUTHORS: Yamano, Shigeyuki
AUTHORS: Izawa, Yoko
AUTHORS: Nakamura, Katsumi
AUTHORS: Harashima, Keiji
TITLE: Elucidation of the Erwinia uredovora Carotenoid Biosynthesis Path
TITLE: by Functional Analysis of Gene Products Expressed in Escheric
JOURNAL: Journal of Bacteriology
VOLUME: 172
ISSUE: 12
PAGES: 6704-6712
DATE: DEC-1990
US-07-783-705A-13

Query Match 37.5%; Score 2624.6; DB 1; Length 5918;
Best Local Similarity 64.7%; Pred. No. 0;
Matches 4007; Conservative 0; Mismatches 2169; Indels 21; Gaps 6;

QY 400 CTTGAACATTTACTGCTCGCGGACAGCAAGCGATCGCGTGGCGTGGCGCGATGCGTGC 459
DB 303 CTTGATCAGTTATTCCTCGTGAGGAGGACGCGATGTTGGTGGCGCGATGCGTGA 362
QY 460 GGAACGCTGGCGGAGGCAACGATATTCGTCCTTTTATTAATGCTGCTGGCAGCGCGGAT 519
DB 363 GGTGCGCTGGCAGCGGAAACGATATTCGCCCATGTTGCTGTTGCTGACCGCGCGGAT 422
QY 520 ATGGTTGCGAGCTGACGCAAAATGGCGTTCTCGATCTCCCTGTGCGATGGAATGGTG 579
DB 423 CTGGGTTGCGCTGTCAGCCATGACGATTAATGATTTGGCTGTGTGGTGGAAATGGTC 482
QY 580 CACCGCGCATCGCTGATTTCTGGATGACATTCCTCGATGGATAACGCGCAGATGCGTCT 639
DB 483 CACGCGCTTCGCTGATCCTTTCAGATATGCTTGCATGACGATGCAAGCTGCGCGC 542
QY 640 GGTGCGCTACCGTGATCGGAATTTGGTGAACACGTGGCGGATTTCTCGCGCCATCGCG 699

```

Db 1618 CTCGGAAGCTTATGCGGCCAGTGAAGAAATTTATGACTGGCTAATGCTGCTCATGACC 1677  
Qy 1780 ACGTGATCCTCAAAACACGCGCGGCGTTTAAATTTGACGAGCGCGCGGATTAATCATCAGT 1839  
Db 1678 GTGTCATTGCCGAACACAGCCACAGAAATGGGCTTAGCGCCCGCCGCAAAAGCTTCAACAGT 1737  
Qy 1840 GCGTCGCGCGTGGCACAATAACAGCCAGATGGTGGCGGCTTTGATTTTTCACGTCAGC 1899  
Db 1738 GTTTTTCGCACTGGCGCAATACAGCAGCTTGTTCCTGAACTGGAATTTTCCCGCAAG 1797  
Qy 1900 AACTGCCGCCCTGCTATACGCGCTGGGCGCACTCGCGCCCGGTTTCTCTGCGCGCGC 1959  
Db 1798 CGTTACCGCTGTTTTCATGCGCTGGGCGCTCTGGCGAAGCAGCACCGTCAACGT 1857  
Qy 1960 TCCATGCGCCCTG---GCCAGCGCTGCTGACGCGGTGTTTATGCTCGCTGGGTAGC 2016  
Db 1858 CTTTCATCCCGTTAATTATCCTCAGAAAAACCCCGGATTTTCGCTCGCTGGGCAAGC 1917  
Qy 2017 TGCAGGCCATCGCTTCCGCGCTTCTGCACTTGGCGCAGCGTGCAGCTGGCGC 2076  
Db 1918 TTACGGGACACCGTTATGGGCTGTTTAAACGATAGTGAAGCCCTGTGAAGAAATTGACG 1977  
Qy 2077 TATCGCTGGTGAATCGCCATTTGGGGGATTAAAGCCGAACAGACGCAATCAGCTGGAGC 2136  
Db 1978 GTACGCTCTGTTAGCCCACTGTGCTGCTTACGGACTCTCAGTGTGAAGACTTGGCGC 2037  
Qy 2137 TCGTGGCGCGCGTGGGTGACGGATTTTCGTGATCAGCGCGAGCCCTACAGCACGCGC 2196  
Db 2038 GAAGCGCTCATACAGAGTGGTGGATTTTGGCCGATCAGTCAGCGCGGCTGTCTCAGGCGC 2097  
Qy 2197 AGCTGTTTATCACTATGCGCGGTTAAACAGCGCGCTGGAAGCACTGGAATCGGTACGC 2256  
Db 2098 AGCTGGCGATCAACACGCGCGCATGAATAACGTTACTGACGCGATTAATTACCGAAGC 2157  
Qy 2257 CGATGCTGGCGTGGCGATTGCTTTTATCAGCGCGGCTGGCGCGCATTTGAGTGGC 2316  
Db 2158 CCTTTTACGCTTCGCTGGCTTTGATCAGCGCGGCTGCGTCAAGCATCGTTTATC 2217  
Qy 2317 ATGAGCTGGTTCGCGCGCATCACGCTTTAGCGGTTTTCATCAACTGGAGCAGCATCTGC 2376  
Db 2218 ACGGCATCGCAAGCGTGTCTCCGCTTTTACCACGAGCCATGCTTTGGCTGTCAGATGC 2277  
Qy 2377 AACAGTGTGACCGACGATGTTAGCGCTAGGATGTACGGATTCAGGCGAGCTGC 2436  
Db 2278 GTTCATTTGTGACCAACGTCGATTTTACAGCAGCATGGCGAATACTCAGACAGCCCTTC 2337  
Qy 2437 AGCGCGCAGCGGTTGCGAGCTGCGCGCATGCTGACAGAGGCGCTGTGCCAGCAGC 2496  
Db 2338 GTTTGGCAGGGGCACTAGCGCGTCCGATATCAATTGAGCAGGTTATGTC-CAACGGT 2396  
Qy 2497 AAGTGTGCTGGCGGAGCGACCTGATGCGCAACGCAATACGATGTGATTTTGGTGGTGC 2556  
Db 2397 CAGCCTGTCTTAAGTGGGAGCGGCTATGCAACGCGATTTATGATCTGATCTCGTGGGCGC 2456  
Qy 2557 TGGACTGGGAATGGCTTGAATTCGCTGCTGCTGCTCAATTTGACGCGCAACTGAATG 2616  
Db 2457 TGGACTGGGAATGGCTTATCGCCTGCTTTCAGCAGCAGCAACCTGATATCGGTAT 2516  
Qy 2617 CCGTGTGCTGGAGAGCGATGCGCATCCGCGAGCGCAATCATCTGCTGCTTTTCATCACAG 2676  
Db 2517 TTTGCTTATCGACGCGCACCCAGCGCGGCGGGAATCATATGCTGATTTTCAACAGA 2576  
Qy 2677 CGATCTCAGCGCGCAACAACTTCGCTGGCTGCAACCGCTGATTAACGCTGTTGTCAGG 2736  
Db 2577 TGATTTGACTGAGAGCAACATCGTTGGATAGCTCGCTGGTGGTTTCATCACTGGCCCGA 2636  
Qy 2737 TTATCAGGTGCTTTTCCGCTGGCGCGCAATCTGGAAGGGGATTAATGTTCCATCGC 2796  
Db 2637 CTATCAGGTACGCTTTCCACACGCGCTGTAAGCTGAAACAGCGGCTACTTTTGTATTAC 2696  
Qy 2797 ATCAGCGAATTTTGGCCGCCATCTTTACGCGGATGGGTGACGATCTGTGGACAAACAC 2856  
Db 2697 TTCTCAGCGTTTCGCTGAGGTTTTCAGCGACAGTTTGGCCCGCACTTGTGGATGGATAC 2756

Qy 2857 AGCCGTAACAACAGGTAAACCCACGACGATGACGCTGGCGGATGCGCGTGAACCTTCTCTC 2916  
Db 2757 CGCGCTCGCAGAGGTTAAATGCGGAATCTGTTCCGTTGAAAAAGGGTCAGGTTATCGGTGC 2816  
Qy 2917 GCAAGTGGTGAATGATGCTGGCGGCTGCAGCGAGCGCACATCTGCAGCTGGGTTATCA 2976  
Db 2817 CGCGCGGTGATTTGACGGGCGGGTTATGCGGCAATTCAGCACTGAGGCGTGGCTTCCA 2876  
Qy 2977 GGTGTTTCTTGAACAAGAGTGGCAGCTGGCGCAGCGCACGCGCTGCAGCAGCCGATCCT 3036  
Db 2877 GCGTTTATTTGGCCAGGAATGCGGATTTGAGCCACCGCATGTTTATCGTCTCCCATTA 2936  
Qy 3037 GATGATGCAACCGCTGATTCAGCAAGCGGTTATCGTTTGTCTACAGCTGCGCTCAG 3096  
Db 2937 CATGATGCAACCGCTGATTCAGCAAAATGGTTATCGTTCGTGTACAGCTGCGCTCTC 2996  
Qy 3097 CGCCGATCGGCTTATTTGATTTGAAGTACCAATTCAGTTAACAGCGCGCTGCGCGGAA 3156  
Db 2997 GCGGACCAAGTTGTTAAATTTGAAGACAGCACTATATGATAATGCAATTAGATCCTGA 3056  
Qy 3157 CACCGCTCTGTCAGCACATCGCGCACTATGCCAATCAGCAAGCTTGACGCTGAGTACGCT 3216  
Db 3057 ATGCGCGCGCAAAATATTTGCGACTATGCCGCGCAACAGGCTTGGCAGCTTCAGACACT 3116  
Qy 3217 GCTGCTGAAGACGACGGCATATTAACGATTAACCTGAGCGGCAACATCGATCGATTCTG 3276  
Db 3117 GCTGCGAGAAGAACAGGCGCCTTACCCATTACTCTGTCGGGCAATGCCGAGCATCTG 3176  
Qy 3277 GCAACAGCAGCGCGCCCAAGCTGCGAGCGCTGCGCGCGGCTGTTTTCATGCCACAC 3336  
Db 3177 G---CAGCAGCGCCCTGCGCTGTAGTGAATTAAGTGGCTGCTTCCCTTACCAC 3233  
Qy 3337 CGGTTACTCTTTCGCTGCGCTGCGCTAGCGAGTTGGTAGCAGCGCTGTGCGCCAC 3396  
Db 3234 CGGCTATTCACTGCGCTGGCGTTGCGGTGCGCGCCGCTGAGTGCACTTGTATCTT 3293  
Qy 3397 CGATGCCCTCAGCTCAGCCAAACATATCGAACGTTTGGTCCCTCAGCACTGGCGCAACA 3456  
Db 3294 TACGTCGGCTCAATTCACCATGCGATACGATTTTGGCCGCGAGCGTGGCAGCAGCA 3353  
Qy 3457 GCGATTTTTCGCTCTAAACCGCATGCTGTTTGGCCGCTGAAGCCGACGACGCTG 3516  
Db 3354 GGGCTTTTTCGCACTGCTGAATCGCATGCTGTTTTCGCGGACCGCGCATTCAGCTG 3413  
Qy 3517 GCGCTGATGCAACGTTTTCACGCTCGATCGCTCGGTTTAAATTAGCGCTTTTTCGCCCG 3576  
Db 3414 GCGGTTATGACGCGTTTATGTTTACCTGAAGATTTAAATTGCCCGTTTATGCGGG 3473  
Qy 3577 GCAACTGCGCTGCGCGATTAACCGGATTTCTGTGCGGCAAGCGCGCTGCGCTATCGG 3636  
Db 3474 AAAACTCAGCTGACCGATCGGCTACGTAATTCAGCGGCAAGCGCTGTTCCGCTATT 3533  
Qy 3637 TGAAGCGCTGCGCGCTGTTGAAATTCGTGCAACAGGGAAGAAAAATGAACCCACT 3696  
Db 3534 AGCAGATTGCAAGCCATATGACGACTCATGCTTAAAGAGCGACTATGAAGAACCACT 3593  
Qy 3697 TATGATTTGGCGAGGCTTTTGGCGCTTGGCGCTGCGGATTTGCGCTGCAAGCGCGGGC 3756  
Db 3594 ACGGTAATTTGGTCAGGCTTGGTGGCTTGGCACTGGCAATTCGTTCTCAAGCTGCGGG 3653  
Qy 3757 ATACCAACACCTTATCTGAGCAGCGGCAAAACCGGCGGAGCGCGCTTATGTTGTAG 3816  
Db 3654 ATCCCGCTCTTACTGCTTGAACAACGTTGATAAACCGCGGCTGCGGCTTATGTTCTACGAG 3713  
Qy 3817 GACAGTGGCTTACCTTCGATGCGGACCCAGCTGATCACCGATCCCGACCGCCATCGAA 3876  
Db 3714 GATCAGGGGTTTACCTTTGATGCGGCGGACGGTTATCACCGATCCCGATGCGCATTA 3773  
Qy 3877 GAGTTGTTTCAGCTGGCAGGAAAAATCGCTCAGCGATTACGCTGAGCTGATGCGGTAACG 3936  
Db 3774 GAACTGTTTGCACTGGCAGGAAAAACAGTTAAAGAGTATGTCGAACTGCTGCGGTTACG 3933

QY	3937	CCCTTCTATCGCTGTGCGAAGATGCGAAACAGCTTGATTACGACAAATATCAGCGC	3996	Db	4914	AGCGCTGTTTTCGGCGCATAAACCGGATATAAAACCATTAATCTCTACCTGGTCGCG	4973
Db	3934	CGGTTTACCGCTGTGTGGAGTTCAGGGAAGTCTTTAATATACGATAACGATCAAAACC	3993	QY	5077	GCCTGACCAATCCAGCGCGCGGTGCGCGGTGATCGGTTCGGCCCAAGCCACCGCC	5136
QY	3997	CTGCTGGAGCAGCAGATCCGACGTTCAATCCGCAAGATGTAGAAGGCTATCGTCAATTT	4056	Db	4974	GCAGGACGATATCCCGCGGAGGATCTCTGGGTCTATCGGTTCGGCAAAAGGACAGCA	5033
Db	3894	CGGCTCGAAGCCGATTCAGCAGTTTAATCCCGCGATGTGCAAGGTTATCGTCAGTTT	3953	QY	5137	AGCTGATGTGAGGATCGCGCGAATGAATGACAGCCTTTATCTTGAAGCAAGTAACGC	5196
QY	4057	CTTGCTATTCACGTGAAGTATTTAGAGGGTTATCTGAACCTCGGACCGGTGCGGTTT	4116	Db	5034	GGTTGATCTGGAGATC--TGAATTAATCCGTCGTTACTCAATCATCGGTCG	5090
Db	3954	CTGGACTATTCACGCGCGGTGTTAAAGAGGCTATCTAAAGCTCGGTACTGTCCCTTTT	4013	QY	5197	AAACCATGGGGTGGCTCGAAGATTTCCGACCCGCGCCAGCTGTTGATGACCCGA	5256
QY	4117	CTGAGGTGCTGACATGCTGCGCTCGGCGCAGTTGGGACGTCGCAAGCATGGCG	4176	Db	5091	AAACGATGGCAGTTGGCTCGAAAGTTTTCGCAAGCTCAAAAGTTATTGATGCAAAA	5150
Db	4014	TTATCGTTTCAGAGACATGCTTCGCGCGCACCTCAACTGGCGAACTCGAGCATGGAGA	4073	QY	5257	CGCGCGCAGCAGCTGATGCTGATGCTGATGCTGCTGCTCACTGCGATGATGTAATGATG	5316
QY	4177	AGCGTCTACAGCATGGTGGCGAAATTTATTCAGGACGATCATCTCGCTCAGCGGTTTCC	4236	Db	5151	CCCGCGCAGCGTACTGATGCTCTACGCTGGTGGCGCCATTGTGACGATGTTATTGACG	5210
Db	4074	AGGTTTACAGTAAAGTTGCCAGTTACATCGAAGATGAACATCTCGCGCAGCGTTTCT	4133	QY	5317	GGCAAAACGCTGGCGAAGCGGCAACGAGATGCGCTCGAAGACGGCAGGCACTATGC	5376
QY	4237	TTCCACTCATGCTGGTGGCGGTAAATCCTTTTGCAAAGTCACTCATCTATACCTTAAT	4296	Db	5211	ATCAGACGCTGGCTTTTCAGGCGCGGCGAGCTGCTCTTACAAACGCGCGCAACGCTGA	5270
Db	4134	TTCCACTCGCTGTGGTGGCGGCAATCCCTTCGCCACCTCATCTCAATTTATACGTTGATA	4193	QY	5377	AGCATCTGCAAAATTTGAACCCGCGGCTTACAGCGCGCGCACATGATGAAACCGCGCT	5436
QY	4297	CATGCGCTGAGCGTGAATGGGCGGTGCTGTTTTCGCGCGCGCACCGCGCGCTGGTG	4356	Db	5271	TGCAACTTGAAGTGAAGAACCGCGCACGCTATGCGAGATGCGAGATGCGACAAACCGCGT	5330
Db	4194	CAGCGCTGAGCGTGAAGTGGGCGTCTGTTTTCGCGTGGCGGCGCACCGCGCATTAAGTT	4253	QY	5437	TTAGGCGGTTTCAGAAAGTGGCGATCATTCACAGTGTGCGCGCAACAACTGCGGTTGATC	5496
QY	4357	CAGGCGATGCGCGACTGTTTCAGAGACTTTGGCGGCGAGCTGTTACTGAATCCGAAGTG	4416	Db	5331	TTGCGGCTTTTCAGAAAGTGGCTATGGCTCATGATATCGCCCGGCTTACGCGTTGATC	5390
Db	4254	CAGGGGATGATAAAGCTGTTTCAGGATCTGGGTGGCGAAGTGGTGTAAACGCCAGAGTC	4313	QY	5497	ATCTGGAAGGCTTCGCTATGGATGTCAGCAACGCAATTAACGGAGCTTCGATGACACGC	5556
QY	4417	AGCAGCTGGAACACAGCGGCATTCGATAGCGGCTTCAGTTAGAGGCGGACGACGC	4476	Db	5391	ATCTGGAAGGCTTCGCGCATGGATGTACGGAAGCGCAATACAGCAACTGGATGATCGC	5450
Db	4314	AGCCATATGGAACAGCAGGAAACAAAGATTGAAGCCGTGCATTTAGAGAACGCTGCAGG	4373	QY	5557	TGCGTTACTGCTATCAGCTCGCGGCGTGGTTCGGTTTGAATGATGGCGCGGTAAATGGCG	5616
QY	4477	TTCGATGCCCGCTGTGGCTCCAAATGCGAGCTGGTGTGCATACCTTACGAAACTGCTT	4536	Db	5451	TGCGCTATTGCTATCAGTTGACAGCGTTGTCGCTTGAATGATGGCGCAATCATGGCGC	5510
Db	4374	TTCTGACGCAAGCGCTGCGCTCAATGCAGATGTGTTTCACTACTATCGCGACCTGTTA	4433	QY	5617	TGCGCGACGAAGCGGTGCTCGATCAGCTCGGATTTAGGACTTGGCGTTCAGCTCACTA	5676
QY	4537	CGCACCATCCGCTGGCAATGAAGCTGACATCCGTGAAGGTTAAGCGATGACGACAC	4596	Db	5511	TGCGGGATAACGCGCACGCTGACCGCGCTGTGACCTTGGGCTGGCATTTTCACTTGAACA	5570
Db	4434	AGCCAGCACCTGCGCGGTAAAGCAGTCCAAACAACTCGAGACTAAGCGCATAGATAAC	4493	QY	5677	ACATTTGCGCGGACATTTAGAAAGATGCGAAATGCTGCTGTATCTGCGCGCAATCCT	5736
QY	4597	TGCTGTTTGTACTTATTTTGGCTTGAATCAGCGCATGAAAGCTCGCGCACCAACACC	4656	Db	5571	ATATTGCTCGCGATATTGTGACGATGCGCATGCGGCGCGCTGTATCTGCGCGCAAGCT	5630
Db	4494	TCTGTGTTGTGCTCTATTTTGGTTGATTCACCATCATGATCAGCTCGCGCATCACAG	4553	QY	5737	GGCTCGATCAGCGCGGATTAACGCGCGGATACGCTGACCTGACCGCAACATCGTCGACGC	5796
QY	4657	GTCTGTTTGGCCCGCTTATCGTGTGATGATGATGATGATTTTCAACAGCAGCCAGCTG	4716	Db	5631	GGCTGAGCATGAAGGTTCTGAACAAAGAGAATTTATGCGCGCACCTGAAACCGTCAGGCGC	5690
Db	4554	GTGTTTTCGSCCGCGTTACCGCGAGCTGATTGACGAAATTTTAAATCATGATGGCCCTC	4613	QY	5797	TGCGCTCACTGGCAGCGGCTTTAGTGGCGGAGGGAACCTTATTTATCACTCGGCGCGAT	5856
QY	4717	GCAGACGATTTTCACTTACCTGCACGCGCCCTGCAGACGCTCGTGGTGGCAACCG	4776	Db	5691	TGAGCCGCTATCGCCCGCTGCTTGGTGGAGAGACGAACTTACTATTGTTGCTGCGCACAG	5750
Db	4614	GCAGAGACTTCTCACATTTATCTGCACGCGCCCTGTGTACGATTCGTCACTTGGCGCT	4673	QY	5857	CCGCTTTACCGGTTTACCGCTGCGCTCGGCGTGGGCCATCGCTACGCTCGCGCGGTTT	5916
QY	4777	CCCGCTGCGCAGCTTTATGTTTATAGCGCGGTGCGCATCTCGGCAACGCTGACATC	4836	Db	5751	CCGCGCTTGCAGGTTGCCCCCTGCTGCTCGCTGGGCAATCGCTACGCGCAAGCAGGTTT	5810
Db	4674	GAAAGTTGCGGAGTTACTATGTGTGGCGCGGTGCGCATTTTAGGCACCGCGAACCCTC	4733	QY	5917	ATCGGAAATTTGGCTCAAAAGTTTACGACGCGCGGTGTCAGCGCTGGGATTCACGCGCAG	5976
QY	4837	GACTGGCAACAGGAAGAACCGCGCTGGCGGATCGAATTTTGGCTTATCTGAGGACGAC	4896	Db	5811	ACCGGAAATATAGTGTCAAAAGTTGAACAGGCGCGCTCAGCAAGCTTGGGATCAGCGGCACT	5870
Db	4734	GACTGGAACGTTGAGGGGCGCAAACTACGCGACCGTATTTTGGTACCTTTGAGCAGCAT	4793	QY	5977	GCACAGTAAAGTGAAGAACTTGGCGCTGCTGGTGAAGGGGCGAGGTTTGGCGATCACTT	6036
QY	4897	TACATCGCGGATTAACGTACCAATTAGTGAACACAGAAATGTTTACGCGGTTTGATTTT	4956	Db	5871	CAACGACCAACCGCGGAAATTAACGCTGCTGCTGCGCGCTCTGTTGAGGCGGCTTACTT	5930
Db	4794	TACATGCTGCTTACGAGTACGCTGTGTCAGCGACCGGATGTTTACGCGGTTTGATTTT	4853	QY	6037	CGCGTGTGTCTGCTGCTGAACCGCGCTCGGCTGCTGTGAGCAGGCTCTCTGTTGATTTT	6096
QY	4957	CGCGACAGCTGCATGCCCATACGCGCTGCGGCTTTGCTGAGGAGCGGATTTTGAACGCA	5016	Db	5931	CCCGGATGCGGCTCATCTCCCGCGCTGCGCATCTCTGGCAGCGCGCGC-----TCT	5984
Db	4854	CGCGACAGCTTAATGCGCTATCATGGCTCAGCGCTTTTCTGTGGAGCGCGTCTTACCCAG	4913	QY	6097	ACGTCGTTGACGCTGCGCAGCGTGGCTTGAAGCTTATTTACGCGGTGGCGCGCTAGAGGAA	6156
QY	5017	AGCGCTGTTCCGCGCGCATAACCGGATATCCGATATCAGCAATCTCTATCTGGTGGGT	5076				



Db 5985 AGGCCATGCTTTCCGGAGCGTCCCTGAAGTTTGTACAGGGCGGCATAGAGAA 6044  
Qy 6157 ACCAAACGACAGCGCTTACGCCCGCGCACCGCATGATGATGCGTGGCCCATGTA 6216  
Db 6045 GCCAAAGAAACACACACCTTCTTTGGCCCTGACGGCGTGATCATACGGTGGCCATATA 6104  
Qy 6217 TAAGCGCTTAAGATAGCTTTCCGGCGGATATAGCGGAACGCCAGCGTTGATGACACAG 6276  
Db 6105 CAACCGTTTGAGTAGCCCTTGGGTGGAATATAGCGGAATGCGCAACGGTTGATGACACAG 6164  
Qy 6277 GCCATCGTCACCATGAAGTAGAGCGCGCTACGTCGTCAATCCGGCACCAATCCACTG 6336  
Db 6165 CCGTGTGTCACCAATAAATAGATTAATCCATAGCGCGTCACTCTGCGCCATCCACTG 6224  
Qy 6337 CAGCGGCCCATGCTTTCGCGGATGACACCGACATAAATACAGCAAAATGCGCAGTACCGCAACAC 6396  
Db 6225 GAGCGGCCCATCTTCTGTACTGCCAGATAAATACAGAGATGCGATAATGCGCAAAAC 6284  
Qy 6397 CACCGCATAAAGATCGTTGAGCTCAAACTTACCGCTGTGGGTTCATGCTGCGACAGATG 6456  
Db 6285 CACGGCATAAAGATCGTTAACTTCAACCGCACCTTTACCGCGTTTCAATGATGTGAAGATG 6344  
Qy 6457 CAGCGCCCATCCCAACCGTGCATGATGATTAATATGCGACAGCGCGCTACGATTTCCAT 6516  
Db 6345 CCATCCCAACCCAGCGCTGCATGATGATTTGTGTGCGAGTGCAGCAATCACTTCCAT 6404  
Qy 6517 CACCAACCGTTGCGCAACAGATAAGCAGCTTCCATACCAAGAGCATTTGCTCCATT 6576  
Db 6405 GCCAATCAGGTAAACGAAACGATCAGGCGATTCCAAATCCCAACATAATTTCTCCGT 6464  
Qy 6577 TGTGNAAGGAGTA 6593  
Db 6465 AGAGACGTCTGCGAGCA 6481

## RESULT 2

US-08-980-832-1/c  
; Sequence 1, Application US/08980832B  
; Patent No. 6291204  
; GENERAL INFORMATION:  
; APPLICANT: Pasamontes, Luis  
; APPLICANT: Tsygankov, Yuri  
; TITLE OF INVENTION: Improved Fermentative Carotenoid Production  
; FILE REFERENCE: Improved Fermentative Carotenoid  
; CURRENT APPLICATION NUMBER: US/08/980,832B  
; CURRENT FILING DATE: 1997-12-01  
; NUMBER OF SEQ ID NOS: 66  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 8625  
; TYPE: DNA  
; ORGANISM: Flavobacterium sp. R1534  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (8348)..(8349)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (8539)..(8540)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (8581)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (8590)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (8592)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (8604)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (8602)..(8604)  
US-08-980-832-1

Query Match

14.8%; Score 1035.4; DB 3; Length 8625;

Best Local Similarity 56.5%; Pred. No. 2.6e-273;  
Matches 2037; Conservative 0; Mismatches 1536; Indels 30; Gaps 5;  
Qy 2486 GTGCAGCAGCAAGTCTGCTGCGGAGGCGACCTGATCGCACGCAATACGATGTGATT 2545  
Db 6984 GGGCGCGCTGCTGAAGGACCGCGAAGGGCGGATCGCAATACATAGCCATGATCTGCTG 6925  
Qy 2546 TTGGTCGGTGTGAGACTGCGCAATGCGTTGATTTGCGTCTGCGTCTGCGTCAATTTGACGCCA 2605  
Db 6924 ATCGGGGCGCGGGGTGTCGGTGTGCGTGTGCGTGTGCGTGTGCGGACCGCAGACCG 6865  
Qy 2606 CAACTGAATGCTGTTGCTGAGAGCGATGCGCATCCGGCAGGCAATACATCTGCTGCG 2665  
Db 6864 GATGCGCGATCGTGAATCTCGACGGCGGTTCGGCCCTTCGACACGACACACTGCTTC 6805  
Qy 2665 TTTCATCAAGCGATCTCAGCGCCGAAACAACTTTCGCTGCTGCAACCGCTGATTACCGTG 2725  
Db 6804 TGCCACGACACGATCTTTCGCCCGCAATGCTGCGCGCTGTCGCCCATTCGTCGCGGC 6745  
Qy 2726 CTTGTCAGGTTATCAGGTGCTTTTCTGCTGCTGCGCGCAATCTGACGGGATTTAT 2785  
Db 6744 GAATGACGGAATCAGAGGTGCTTTCCGACCATTCGCGCGCTGACGACAGGCTAT 6685  
Qy 2786 TGTTCATCGCATCAGGCGATTTTTCGCCCATCTTTACGCGCGATGCGTGTGACGATCTG 2845  
Db 6684 GGTCTGATCGAGCGCGCGCTGATCGGGCTGTGCA-----GGGTGTCGATCTG 6634  
Qy 2846 TGGCAAAACACAGCGGTCAACAGGTAACACACGAGGTGACGCTGCGGATGCGCGT 2905  
Db 6633 CGGTGGAATACGATGTCGCGACGCTTGGACGATACCGCGCGACGCTGACGAGCGCTG 6574  
Qy 2906 GAACTTGTGCGCAAGTGTGATTTGATGTCGGGCTGCGACGCGCAATCTGTCAG 2965  
Db 6573 CGGATCGAGGCTGCTTTCGCGCATGACGCGCGTGTGCGCTGAGACCGCGACCTGACC 6514  
Qy 2966 CTGGGTTATCAGTGTTCCTTGGACAAGATGCGAGCTGCGCGACGCGCGCTGCGAG 3025  
Db 6513 GTGGGTTTCAGAAATTCGTGGCGGTGAGATCGAGACCGACGCGCGCGCTGCGAG 6454  
Qy 3026 CAGCGGATCTGATGATGCCACCGTCAATCAGCAAGCGGTTATCGTTTGTCTACAG 3085  
Db 6453 CGCGCGATGATCATGAGCGGACCGTTCCGCAAGTGAAGGTTACCGCTTCTATCTG 6394  
Qy 3086 TGTGCGCTCAGCGCGATGCGCTATTGATTTGAAGTACCAATTCATGTTACGACGCGCG 3145  
Db 6393 CTGCGCTTCAATGCCACCGCATCTGATCGAGGATACGCTTACAGCAGCGCGCGAT 6334  
Qy 3146 CTGGCGGAGAACACCGCTGTCAGCACATCGCGGCTATGCCAATCAGCAAGGCTGAGC 3205  
Db 6333 CTGGACGATGGCGCGTGGCGCAGGGTCTGTTGACTATGCCCGCAGCGCGGCTGAGC 6274  
Qy 3206 CTGAGTACGCTGCTGTAAGAGCAGCGCATATTACCGATTACCTGAGCGGCAACATC 3265  
Db 6273 GGGCAGGATGCGGCGCGAAGGGGCTCTCTGCCCATCGCTGCGCGCATGACGCATA 6214  
Qy 3266 GATCGATTCTGGCAACAGCAGCGCGGCAAGCGTGCAGCGGCTGCGCGCGGCTGTTT 3325  
Db 6213 GGTCTTCTGGCGGACCAACGCGGAGGGGCGTCCGTTTGGGCTGGGGGAGGGCTGTTT 6154  
Qy 3326 CATGCCACCGGTTACTCTTTCGCTCGCGGTGGCGCTAGCGAGTTGGTAGGAGCG 3385  
Db 6153 CACCGCGTCAACCGGATTTTCGTGCTTATGCCGCGCATGTCGCGGATGCCATTCGCGGCG 6094  
Qy 3386 CTGTTGCCCAACCGATGCCCTCAGCGTCAAGCAACATATCGAAGCGCTTTGCGCGCTCAGCAG 3445  
Db 6093 CGGACCTGACGACCGGCTCCCGCTCGCGGGTGGCGGCTGGCCATCGATCGCGCG 6034  
Qy 3446 TGGCGCAACAGCGATTTTTCGTCTGCTAAACCGCATGCTGTTTTTGGCCGCTAAGCCG 3505  
Db 6033 GATCGCG---ACCGCTTCTGCGGCTGCTGTAACCGGATGCTGTTCCGCGGCTGCCGCC 5977  
Qy 3506 CAGCAGCGCTGGCGGCTGATGCAAGTTTTCACGGCTCGATGTCGCGGTTAATTAGCCGC 3565



QY 5726 GCGCAATCTGCTCGATCAGCGGGATTACGCGCGGATACGCTGACTGACCGCAACA 5785  
Db 3765 GCTTCCGACTGCTGCGCGAGCGGG-----GCGACGGTTGAGGGTCCGGTGCC 3715  
QY 5786 TCGTGCAGCGCTCGCTCACTCGCAGCGGCTTTAGTGGCGGAGCGGAACCTATTATCA 5845  
Db 3714 TTCGACGGCTCTATTTCGTCATCATCGCTGCTTGACGGCGGAGCCCTATTATGC 3655  
QY 5846 CTCGCGCGATCCGGTTTACCGGTTTACCGTGGCTCGCGTGGGCCATCGTACGGC 5905  
Db 3654 CTCGCGCGGAGGCGCTTCCGATCTGCGCGCGCTGCGGCTGCTCGATCGCGCGCG 3595  
QY 5906 TCGCGCGTTTATCGGAATTTGGCTCAAAGTTGACGACGCGGCTGTGCAAGCTTGG 5965  
Db 3594 GCTCGTATCTATCGCGCAATCGGACGCGCATCGCGAGGTTGCGCCCGAGGCTTATCG 3535  
QY 5966 TTCAGCGGACGCGCAGTAAAGTTGAAAACTGGCGCTGCTGGTGAAGGGGCGAGTTT 6025  
Db 3534 CACGCGGATCAGACGCTCGAAGCTGCCAAGATCGGGCTTCTGGCGCGGAGGCTTGA 3475  
QY 6026 GCGCATCACTTCGCGTGTGCTGCTCTGAACCGGCTCGCGTGGTCTGTGCAAGCTCC 6085  
Db 3474 CGCGCGCATCGCGCTGCGCGCGCGGAATCAGCGCGAGCGCTGTGCAAGCGGACC 3415  
QY 6086 TCG 6088  
Db 3414 GCG 3412

## RESULT 3

US-09-920-923B-1/c

; Sequence 1, Application US/09920923B

; Patent No. 6671134

; GENERAL INFORMATION:

; APPLICANT: Pasamontes, Luis

; APPLICANT: Tevankov, Yuri

; TITLE OF INVENTION: Fermentative Carotenoid Production

; FILE REFERENCE: 15464 US (C38435/125944)

; CURRENT APPLICATION NUMBER: US/09/920,923B

; CURRENT FILING DATE: 2001-08-02

; PRIOR APPLICATION NUMBER: 08/980,832

; PRIOR FILING DATE: 1997-12-01

; NUMBER OF SEQ ID NOS: 66

; SOFTWARE: Patent in version 3.1

; SEQ ID NO 1

; LENGTH: 8625

; TYPE: DNA

; ORGANISM: Flavobacterium sp. R1534

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (8348)..(8349)

; OTHER INFORMATION: unsure

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (8539)..(8540)

; OTHER INFORMATION: unsure

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (8581)..(8581)

; OTHER INFORMATION: unsure

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (8590)..(8590)

; OTHER INFORMATION: unsure

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (8592)..(8592)

; OTHER INFORMATION: unsure

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (8602)..(8604)

; OTHER INFORMATION: unsure

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (8602)..(8604)

; OTHER INFORMATION: unsure

US-09-920-923B-1

Query Match 14.8%; Score 1035.4; DB 4; Length 8625;  
Best Local Similarity 56.5%; Pred. No. 2.6e-273;  
Matches 2037; Conservative 0; Mismatches 1536; Indels 30; Gaps 5;  
QY 2486 GTGCGAGCAGCAAGTCTGCTGCGGAGGCGAGCTGATCGCGACGCAATACGATGTGATT 2545  
Db 6984 GGGCGCGCTGCTGAAGGACCGGAGGGCGGATCGCAATACATGAGCCATGATGTGCTG 6925  
QY 2546 TTGCTCGTGTGACTGCGCAATGCTTTGATTTGGCTGCTGCTGCTGCTGCTGCTGCTG 2605  
Db 6924 ATCGCGGCGCGGGCTGTCCGCTGCTGATCGCGCTTGCCTTGCCTGCGACCGACGCG 6865  
QY 2606 CAACTGAATGCTTGTGAGAGCGATGCGCATTCGCGCAGGCGAATCATACCTGCTG 2665  
Db 6864 GATGCGCGCATGCTGATGCTGACGCGCGGTCCGCGCCCTCGGACCGACACACCTGCTC 6805  
QY 2666 TTTCATACAGCGATCTAGCGCGCAACAACTTCGCTGCTGCTGCTGCTGCTGCTGCTG 2725  
Db 6804 TGCCACGACACGATCTTTGCGCCGAAATGCTGCGCGCTTGTGCGCCATTCGTCGCGG 6745  
QY 2726 CGTTGCTCAGGTATCAGGTGCTTTTCTGCGCTGCGCGCAATCTGACCGGATTTAT 2785  
Db 6744 GAATGACGATCAGAGGTCCGCTTCCGACCATTCGCGCGCTGACGACGAGCTAT 6685  
QY 2786 TGTTCATGCTAGGCGATTTTCCCGCCATCTTTACGCGCGATGCGGTGACGATCTG 2845  
Db 6684 GGTGATCGAGCGCGCGCTGATCGGGCTGTGCA-----GGTGTGCTGATCTG 6634  
QY 2846 TGGCAACACAGCGGTACACAGGTAAACCCAGCGAGTACGCTGCGGATGCGCGT 2905  
Db 6633 CGTGGAAATACGATGTGCGACGCTGGAAGATACCGCGCGCATGACGCGAGCGCTCG 6574  
QY 2906 GAACTTGTGCGCAAGTGTGATTTGATGTGCGCGCTGCGACCGCGCACATCTGCAG 2965  
Db 6573 CGGATCGAGGCTGCTGCTGATCGAGCGCTGCTGCGCTGCGAGACCGCGACCTGACC 6514  
QY 2966 TTGGGTTATCAGGTGTTTCTGGAAGAAGTGGCGAGTGGCGAGCGCGCGCTGCGAG 3025  
Db 6513 GTGGGTTTCCAGAAATTCGTGGCGCTCGAGATCGAGACCGCGCGCGCGCTGCGAG 6454  
QY 3026 CAGCGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3085  
Db 6453 CGCCGATGATCATGAGCGAGCGCTTTCGCGAGATGAGCGGCTGCTGCTGCTGCTGCTG 6394  
QY 3086 CTGCGCTCAGCGCGATCGGCTATTGATTTGAAGATACCATTTACCTTAAACGAGCGCG 3145  
Db 6393 CTGCGCTTCACTCCACCGCATCTGATCGAGGATACCGCTACAGCGCGCGCGGAT 6334  
QY 3146 TTGGCGGAGAACACCGCTGCTGAGCACATCGCGCATATGCGCAATCAGCAGGCTGAGCG 3205  
Db 6333 CTGAGCATGCGCGCTGCGCGAGCGCTGCTGCGATATGCGCGCGCGCGCGCTGCGAGC 6274  
QY 3206 CTGAGTACGCTGCTGCTGAGAGCACCGCATATTACCGATTACCTTACGCGCGCAACATC 3265  
Db 6273 GGGCAGGAGATCGCGGAGGAGGCGCATCTCGCCCATCGCGCTGCGCGCGCGCAT 6214  
QY 3266 GATCGATTCTGCAACAGCAGCGCGCGCAAGCGCTGCGCGCGCTGCGCGCGCGCTGCTT 3325  
Db 6213 GGTCTTCTGCGCGACCAACGCGCGCGGCGGCTGCGCGCTGCGCGCGCGCGCTGCTT 6154  
QY 3326 CATGCCAACCGGTTACTCTTTCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCG 3385  
Db 6153 CACCGCGTCAACCGGATATTCGCTGCGCTATGCGCGCGAGGTGCGCGATGCCATCGCGCG 6094  
QY 3386 CTGTTGCGCGATGCGCTCAGCTCAGCAACATATCGAAGCTTTTCCCGCTGAGCAG 3445  
Db 6093 CGCGACCTGACAGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTG 6034  
QY 3446 TGGCGCAACAGCGATTTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3505  
Db 6033 GATCGCG---ACCGCTTCTGCGGCTGCTGATACCGGATGCTGTTCCGCGCGCTGCCGCC 5977



Db 3825 CCAGCTGACGAACATCGCTCGCGACGTCGATCGACGATCGCCCATCGGGCGCTGCTATCT 3766  
Qy 5726 GCCGCAATCTCGCTCGATCAGCGGGGATTCAGCGCCGATACGCTGACTGCAACGCAACA 5785  
Db 3765 GCCTCGGACTGGCTGGCCGAGCGGG-----GGCGACGGTTGAGGGTCCGGTGCC 3715  
Qy 5786 TCGTGCAGCGCTCGCTCACTGCGACGCGGTTAGTGGCGGAGGCGGAACCTTATATCA 5845  
Db 3714 TTGCGACGCGCTCTATTCGGTCATCATCGCTGCTTGACGCGCGCGAGCCCTATTTATGC 3655  
Qy 5846 CTCGCGCGCATCCGGTTTACCGGGTTTACCGCTGCGCTCGCGTGGGCCATCGCTACGCG 5905  
Db 3654 CTCGCGCGCGAGGGCTTCGCACTGCGCGCGCTCGCGTGGTTCGATCGCGCGCG 3595  
Qy 5906 TCGCGCGGTTTATCGCGAAATTTGGGCTCAAGTTTCAGCACGCGCGTGTGCAACGCTTGGGA 5965  
Db 3594 GCTGGTATCTATCGCGCAATCGGACGCGCATCGGACGGGTGCGCGAGGCTATCG 3535  
Qy 5966 TTCAGCGGACGACCAAGTAAAGGTGAAAACTGGCGCTGCTGGTGAAGGGCGAGGTTT 6025  
Db 3534 CAGCGGATCAGCAGCTGCAAGGCTGCCAAGATCGGGCTTCTGGCGCGCGAGGCTTGA 3475  
Qy 6026 GCGGATCACTTCGCGTGTCTGCTCTGAACCGGCTCGGCTGGTGTCTGTGCAGCGTCC 6085  
Db 3474 CCGCGCGCATCGCGCTCGCGCGCGCGGCGGAATCAGCCGCGACGCTGTGACCGGACC 3415  
Qy 6086 TCG 6088  
Db 3414 GCG 3412

RESULT 4

US-08-980-832-27/c  
; Sequence 27, Application US/08980832B  
; Patent No. 6291204  
; GENERAL INFORMATION:  
; APPLICANT: Pasamontes, Luis  
; APPLICANT: Tsygankov, Yuri  
; TITLE OF INVENTION: Improved Fermentative Carotenoid Production  
; FILE REFERENCE: Improved Fermentative Carotenoid  
; CURRENT APPLICATION NUMBER: US/08/980,832B  
; CURRENT FILING DATE: 1997-12-01  
; NUMBER OF SEQ ID NOS: 66  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 27  
; LENGTH: 11233  
; TYPE: DNA  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism: Plasmid pZea4  
US-08-980-832-27

Query Match 14.8%; Score 1035.4; DB 3; Length 11233;  
Best Local Similarity 56.5%; Pred. No. 3.1e-273;  
Matches 2037; Conservative 0; Mismatches 1536; Indels 30; Gaps 5;  
Qy 2486 GTGCCAGCAAGTGTCTGTGCGGAGCGACCTGATCGCACGCAATACGATGTGATT 2545  
Db 7666 GGGCGCGCTGCAAGGACCGGAAGGGCGGATGCAATACATGAGCCATGATCTGCTG 7607  
Qy 2546 TTGGTCGGTGTGACTGCGGAATGGCTTGATGTGGCTGCTGCTGCTGCTGCTGCTG 2605  
Db 7606 ATCGGGGCGGGGCTTCCGGTGGCTGATCGCGCTTTCGGCTTTCGGACCGGACGCG 7547  
Qy 2606 CAACTGAATGCTGTGCTGGAGAGCGATGCGATCCGGCAGGCAATACATCTGCTGCG 2665  
Db 7546 GATCGCGCATCGTATGCTGACGCGCGGTTCGGCCCTTCGACACGACACCTGTGCTC 7487  
Qy 2666 TTTCATCAGCGGATCTCAGCGCGGCAACTTTCGCTGGCTGCAACCGCTGATACCGGTG 2725  
Db 7486 TGCCACGACAGGATCTTTCGCCCGAATGGCTGGCGCGCTGTGCGCCCATTCGTCGGCG 7427  
Qy 2726 CGTGTGTCAGGTATCAGGTGCGTGTTCCTGCGCTCGCGCGCAATCTGACGGGATAT 2785

Db 7426 GAATGGAAGGATCAGAGGTCCGCTTCCCGACCATTCGCGCGCTGACGACAGGCTAT 7367  
Qy 2786 TGTTCATTCGATCAGGGGATTTTCCCGGCATCTTTTACGCGCGATGGGTGACGATCTG 2845  
Db 7366 GGTTCGATCGAGGCGGCGCTGATCGGCTGCTGCA-----GGGTGCTGATCTG 7316  
Qy 2846 TGGCAAAACACAGCGCTCAACAGAGTAAACCCACGAGGTGACGCTGCGGATGCGCGT 2905  
Db 7315 CGGTGGAATACGATGTCGACGCTGAGACGATACCGGCGGACGCTGACGAGCGCTG 7256  
Qy 2906 GAACTTGTGCGCAAGTGTGATTTGATGTCGCGGCTGCGACCGGACGACATCTGCGAG 2965  
Db 7255 CGGATCGAGGCTGCTGCTGATGACGCCCTGTGTGCGCTCGAGACCCCGACCTGACC 7196  
Qy 2966 CTGGGTTTATCAGGTGTTCTTGGCAAGAGTGGCAGCTGGCGCAGCGCACGCTGCGAG 3025  
Db 7195 GTGGGTTCAGAAATTCGTGGCGCTCGAGATCGAGACCGACGCCCTCATCGGCTGAG 7136  
Qy 3026 CAGCGGATCTGATGATGATGCAACCGTTCGATCAGAGAGCGGGTTATGCTTTGTCTACAG 3085  
Db 7135 CGCCCGATGATCATGAGACGCGACCGTTCCGACAGATGGAACGGGTACCGCTTCTATCT 7076  
Qy 3086 CTGCGCTCAGCGCGGATCGGCTATTGATTGAAGATACCATTAAGTTAACGACCGCGG 3145  
Db 7075 CTGCCCTTTCAGTCCACCGCATCTGATCGAGGATACCGCTACAGCAGCGCGCGAT 7016  
Qy 3146 CTGGCGGAGAACACCGCTCTGTCAGCACATCGCGACTATGCCAATCAGCAAGGCTGGAG 3205  
Db 7015 CTGGACGATGGCGCTGGCGCAGGCTGCTGGACTATGCCCGCAGCGGGGCTGAGCC 6956  
Qy 3206 CTGAGTACGCTGCTGCTGAGAGACACGGCATATTACCGATTACCTTGAGCGGCAACATC 3265  
Db 6955 GGGCAGGAGATCGCGCGGCAAGGGGCTATCTCGCCATCGCGCTGCGCCATGACGCCATA 6896  
Qy 3266 GATCGATTCTGCAACACGCGCGCGCAAGCGTGCAGCGGCTGCGCGCGGCTGTTT 3325  
Db 6895 GGTCTTCTGGCGGACACGCGCAGGGGGCGGTGCGGTTGGGCTGGGGCAGGGCTGTTT 6836  
Qy 3326 CATGCAACACCGGTTACTCTCTTGGCTGCGCGCTGCGGCTAGCGGAGTTGGTAGCAGCG 3385  
Db 6835 CACCCGCTCAGCGGATATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6776  
Qy 3386 CTGTTGCGCACCGATGCTGCTCAGCTCAGCGCAACATATCGAACGCTTTTCCGCTGACGAG 3445  
Db 6775 CGCGACCTGACGACGCGCTGCGCGCTGCGCGGCTGCGGCTGCGGCTGCGGCTGCGG 6716  
Qy 3446 TGGCGGACAGCGATTTTCCGCTGCTGCTTAACCGCATGCTGCTTTTGGCGCGTAAAGCG 3505  
Db 6715 GATCGCG---ACCGCTTCTGCGGCTGCTGAAACCGGATGCTGTCGCGGCTGCCCGCCC 6659  
Qy 3506 CAGCAGCGCTGGCGGCTGATGCAAGCTTTTATCCGCGCTCGATGCGCGGTTAATTAGCCGC 3565  
Db 6658 GACCGTCTGCTATCGCTGCTGAGCGGTTCTACCGCTGCGCGCAGCGCTGATCGAGCGC 6599  
Qy 3566 TTTTACGCGCGGCAACTGCGCTGCGCGATATAAAGCGGATTTCTGTGCGGCAAGCGCGC 3625  
Db 6598 TTCTATGCGCGGCTGACATTGCGCGACCGGCTTGCATCGTCAACCGGACCGCGCGCC 6539  
Qy 3626 GTGCCCATCGGTGAAGCGCTGCGCGCTGTTGAATTCTGTCGAACACGAGGAGAGAAAA 3685  
Db 6538 ATTCGCTGTTCGACGCGCTGCGCTGCTGCGCGAACGCCCTGCTGCGAGGAGAGCA 6479  
Qy 3686 TGAACCGCATTTATGATTTGGCGCAGGCTTTGGGCGGCTGCGGCTGCGGATTTGCGCTG 3745  
Db 6478 TGAGTTCCGCATCTGATCGGCGGAGTTTGGCGGGCTTGGCTTGCATCCGCTGTC 6419  
Qy 3746 AAGCGCGGCGCATACAACACCTTACTGAGAGAGCGCGCAAAACCGCGCGGACGCGCT 3805  
Db 6418 AATCGCGGCGCATCGCGACCACTGTCGAGCGCGCGCAAGCGCGCGCGCGCGCT 6359  
Qy 3806 ATGCTTTGAGACAGTGGCTTTACTTTCGATGCGCGGACCCACGCTGATCACCGATCCCA 3865



Db 6358 ATGCTGTGAAACGATCAGGCGCCACGCTCTTCGATGAGGCCCGCAGCGGTCTGTGACCGACCCCG 6299  
 Qy 3866 GGGCCATCGAAGAGTTGTTTACGCTGGCAGGAAATCGCTCAGCGATTACGCTCAGAGCTGA 3925  
 Db 6298 ACAGCCTGCGAGAGTGTGGGCGCTTCAGCGGCAACCGATGGAGCGTGAAGCTGC 6239  
 Qy 3926 TCCCGGTAAAGCCCTTCTATCGCCCTGTGCTGGGAAGATGGCAAAACAGCTTGAATACAGACA 3985  
 Db 6238 TCCCGGTCTCGCCCTTCTACCGCTGACATGGGCGGACGCGCGAGCTTCGAATACGTGA 6179  
 Qy 3986 ATAAATCAGCGCTGTGTGAGCAGAGATGCGCAGTTCAATCCGCAAGATGAGAAGCT 4045  
 Db 6178 ACGACGACGACGAGCTGAATCCGCCAGGTGCGCTCTCAATCCGCCGATGTGATGCT 6119  
 Qy 4046 ATGCTCAATTTCTTGCTATTCACTGTAAGTATTTAGAGAGGTTATCTGAACCTTGGCA 4105  
 Db 6118 ATCGCCGCTTCCAGGATACGCGGAGGTCTATCGCGAGGGTATCTGAAGCTTGGGA 6059  
 Qy 4106 CGGTGCGGTTTTCGAGGTGCGTGAATGCTCGGCTGCGCGCGCAGTTGGAGCTGTC 4165  
 Db 6058 CCAAGCCTTCTGAAAGCTGGGCGAGATGCTGAACGCGCGCGGCTGATGCGCCTGC 5999  
 Qy 4166 AAGCAGTGGCAGCGCTACAGCATGTCGCAAAATTTATTCAGGACGATCATCTGCGTC 4225  
 Db 5998 AGGCATACCGCTCGGTCCAGCATGTTGGCGGCTTTCATCCAGGACCGCATCTGCGGC 5939  
 Qy 4226 AGCGGTTTCTTCCACTCATTTGCTGCTGGGCGGTAAATCTTTTGAAGCTCATGATCT 4285  
 Db 5938 AGGCTTCTGTTCCACAGCTGCTGCTGGTGGGGAACCGCTTTTGACAGCTCGATCT 5879  
 Qy 4286 ATAGCTTAATCATGCGCTGAGCGTGAATGGGCGGTGTTTCCGCGCGCGCACCG 4345  
 Db 5878 ATGCGCTGATCCATGGCTGGAACGCGCGCGCGCTGTTCTGGCGGCACTGCTGCTGA 5759  
 Qy 4346 GCGGCTGCTGCAAGGATGCGCGAGCTGTTTCAGAGCATTTGGCGCGGAGCTGTTACTGA 4405  
 Db 5818 ACCAGCTGGTTCGCGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGA 5759  
 Qy 4406 ATGCGGAAGTGAACGCTGGAACAGCGGCAATCGCATTAGCGGCTTCAGTTAGAG 4465  
 Db 5758 ATGCGCGCTCAGCGGATCGACACCGAGGCGATCGCGCACGCGGCTCAGCTGCTGG 5699  
 Qy 4466 GCGGACGAGCGTTGATGCGCGCTGTCCTCAATGCGGAGCTGTTGCTGATACCTACG 4525  
 Db 5698 ACGGCGGCGAGTTGCGCGGATACGTTGCGGAGCAACGCGGAGCTGATGACAGCTATC 5639  
 Qy 4526 ACAAACTGCTTCGCCACCATCCGCTGGCAATGAACCGTGGACATCGCTGAAGGTAAGC 4585  
 Db 5638 GCGACCTGCTGGGCAATACCGCGCGCGGCGCACAAAGCGCGGATCTTGAACCGGAGC 5579  
 Qy 4586 GCATGAGCAACTCGCTGTTGCTACTATTTTGGCTGGAATCAGCCGCAATGAACAGCTCG 4645  
 Db 5578 GCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5519  
 Qy 4646 GCGACCAACGCTGTTTGGCGCGGTTATTCGATGATGATGATGATGATGATGATGATGAT 4705  
 Db 5518 CCCACCAAGCGTCACTTCGCGCGCGCTACAAAGGCGCTGGTGAACGAGATCTTCAACG 5459  
 Qy 4706 GAGCAGCTGCGCAGACGATTTTCACTTACCTGCAAGCGCGCTGCGAGAGCGATCGGT 4765  
 Db 5458 GCGCAGCTGCGGAGCATTTCTCGATGATGATGATGATGATGATGATGATGATGATGAT 5399  
 Qy 4766 CGCTGGCACCGCGCTGCGGAGCTTTTATGTTGTTAGCGCGCGGTGCGCATCTGGCA 4825  
 Db 5398 GCTTGGCCCCGAGGGATGTCCAGCATAGCTTCTTGGCCGCTTCGCGCATCTGGGCGC 5339  
 Qy 4826 CCGCTGACATCGACTGGCAACAGGAGCGCGCTTGGCGGATCGGATTTTCTGTTATC 4885  
 Db 5338 GCGCGCATGCTGATTTGGGAAGCGGAGCGCGGCTATGCGGAGCGCATCTTCGAGGAAC 5279  
 Qy 4886 TGGAGCAGCATACATGCGCGGATTAGCTGACCAATTAGTGACACACAGAAATGTTTACGC 4945  
 Db 5278 TGGAGCGCGCGGCAATCCCGAGCTCGGCAAGCACTTGACCGCTCAGCGCGCATCTTCAGCC 5219

Qy 4946 CGTTTGTATTTTCGCGACACGCTGTCATGCCATCAGGCTCGCGGTTTTCGCTGGAGCGCA 5005  
 Db 5218 CGCGCATTTTCAGCACCGACCTGTGCGGCCATCAGCGAGCGCTTCTCGTGGAGCGCA 5159  
 Qy 5006 TTTTGAACGCAAAAGCGCTGTTTCGCGCGCATAAACGCGATGCCATATCAGCAATCTCT 5065  
 Db 5158 TCTGACGCAATCCGCTGTTTCGCGCGCATAAACGCGACCGCGCATTCGCAACTTCT 5099  
 Qy 5066 ATCTGTTGGTGGCTGCTGCTAGCATCCAGGCGCGGCGTTCGCGGCTGATCGGTTTCGGCA 5125  
 Db 5098 ACATCTGTTGGGCGGCGACGCTCCGCTGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCT 5039  
 Qy 5126 AGGCCACCGCAGGCTGATGCTGGAGGATCGCGCGCAATGAATTCGACAGCTTTTACTGA 5185  
 Db 5038 AGGCCACCGCAGGCTGATGCTGCTGCGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4979  
 Qy 5186 GCAAGTAAACGCAAAACATGCTGGCTGGCTGAGAGTTTCGCCACCGCGCAAGCTGTT 5245  
 Db 4978 TTCCGAAGCG---GCCATCGCGAGGTTTCGAAAGCTTCGCGCAGCGCGCTGAT 4922  
 Qy 5246 TGATGCAACCGACCGCGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5305  
 Db 4921 GCGCGCGGCTGCGCGAGGATAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4862  
 Qy 5306 TGTGATGATGGGCAAAACGCTGGCGGAAAGCGGCGACGAGCATGCGTTCGAAAGCGGCA 5365  
 Db 4861 CGTGATCGACGCGGAGGCTGATGGGTTTCTGCCCGCGAG-----GCGGCGCGCGCACCA 4808  
 Qy 5366 GGCAGCTATGCGCATCTGCAAAATTTGAACCGCGCGCTACAGCGCGCGCGCACATGGA 5425  
 Db 4807 GCGCGGCTGGGCGCTGCGCGCGACGCTGCGCGCTGCTGCGCGCTGCTGCGAGGACGCGCGCT 4748  
 Qy 5426 TGAACCGCGCTTTAGGCGTTTTCAGGAGTGGGCTGCTGCGCGCGCTGCTGCGAGCTTTG 4688  
 Db 4747 GTCGCGCGCTTTCGCGCGCTGCGCGAGCTGCGCGCGCTGCTGCGAGCTTTG 4688  
 Qy 5486 GCGGTTGATCATCTGGAAGCTTTCGCTATGATGATGATGATGATGATGATGATGATGATGAT 5545  
 Db 4687 GCGGATGACCTGATCGAGGTTTTCGCGATGATGATGATGATGATGATGATGATGATGATGAT 4628  
 Qy 5546 CGATGACACGCTGCTTACTGCTATCAGCTGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5605  
 Db 4627 GATGACGCTGCTGGAATTTCTTACCACTGCGCGGCTGCTGCGGCTGCTGATGATGATGATGAT 4568  
 Qy 5606 GCTAATGGGCTGCGCGAGGAGCGGCTGCTGATCAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 5665  
 Db 4567 GGTGATGGGCTGCGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4508  
 Qy 5666 CCAGCTCACTAATTCGCGCGCATTTGTAGAAAGATCCGAAATGCTGCTGCTGCTGCTGCTGCTGCT 5725  
 Db 4507 CCAGCTGAGCAACATCGCTGCGGAGCTGATCGACGATCCGCGCTGCTGCTGCTGCTGCTGCTGCT 4448  
 Qy 5726 GCGCAATCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5785  
 Db 4447 GCTTGGCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4397  
 Qy 5786 TCGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5845  
 Db 4396 TTTGGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4337  
 Qy 5846 CTGCGCGCATCGGTTTACCGGTTTACCGGTTTACCGGTTTACCGGTTTACCGGTTTACCGGTTTACCGG 5905  
 Db 4336 CTGCGCGCGGCGGCTTTCGCGCATCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4277  
 Qy 5906 TCGCGGCTTTATTCGCGAAATTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5965  
 Db 4276 GCT 4217  
 Qy 5966 TTCACGCGCAGCGCACAGTAAAGGTGAAAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6025  
 Db 4216 CCAGCGGATCAGCACGCTCGAGGCTGCCAAGATCGGCTTCTGCGCGCGCGAGGCTTGGGA 4157







NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESS: Ladas & Parry  
STREET: 26 West 61 Street  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10023  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 720Kb storage  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: N/A  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/783,705A  
FILING DATE: 19911023  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 1-103078  
FILING DATE: 21-APR-1989  
APPLICATION NUMBER: JP 2-53225  
FILING DATE: 05-MAR-1990  
APPLICATION NUMBER: US 07/519,011  
FILING DATE: 19-APR-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Schwadron, Janet I.  
REGISTRATION NUMBER: 33,778  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-708-1935  
TELEFAX: 212-246-5959  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1479 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: other nucleic acid (plasmid DNA)  
US-07-783-705A-10

Query Match 11.8%; Score 826.2; DB 1; Length 1479;  
Best Local Similarity 72.6%; Pred. No. 2.5e-216;  
Matches 1068; Conservative 0; Mismatches 403; Indels 0; Gaps 0;

QY 3685 ATGAACGCACTATGTGATGGCGAGCGCTTTGGCGGCTGGCGCTGGCGATTCGCCCTG 3744  
Db 1 ATGAACCACTACGCGTAATGGTGCAGGCTTCGGTGGCTGGCACTGGCAATTCGTCTA 60  
QY 3745 CAAGCGGCGGCATACCAACCACTTACTCGAGCAGCGGACAAACCGGCGGAGCGCC 3804  
Db 61 CAAGCTGCGGGGATCCCGCTTACTGCTTGAAACAGTGATGAAACCGGCGGTCGGGCT 120  
QY 3805 TATGTGTTTGGAGACAGTGGCTTTTACCTTCGATGCGGACCCACCGTGATCCCGATCCC 3864  
Db 121 TATGCTACAGGATCAGGGGTTTACCTTTGATGAGGCGCCGAGGTTATCACCGATCCC 180  
QY 3865 AGCGCATCGAAGTGTTCAGCTGGCAGGAAATGCTCAGCGAATTCGTCGAGCTG 3924  
Db 181 AGTGCATTAAGAACTGTTTTCAGCTGGCAGGAAACAGTAAAGAGTATGTCGAACTG 240  
QY 3925 ATGCGGTAAACCGCTTCTATCGCTGCTGCGAGATGGCAACAGCTTGATACGAC 3984  
Db 241 CTGCGGGTTACGCGGTTTACCGCTGTTGCGGAGTCAGGGAAGGCTCTTAAATACGAT 300  
QY 3985 AATAATCAGCGCTGCTGGAGCAGCAGATCGCCACCTTCAATCCGCAAGATGTAGAAGC 4044  
Db 301 AAGCATCAACCGGCTCGAAGCGCAGATTCAGCAGTTTAAATCCCGCGATGTCGAGGT 360  
QY 4045 TATGTCATTTCTGCTTATTCAGCTGAAGTATTTAGAGAGGTTATCTGAAATCTCGGC 4104  
Db 361 TATGTCAGTTTCTGACTATTCAGCGCGGTTGTTTAAAGAAAGGCTATCTAAAGCTCGT 420  
QY 4105 ACGGTGCCCTTCTGACAGTGGTGACATGCTCGGCTCGCGCGAGTTGGAGCTCTG 4164

RESULT 7  
US-08-095-726-7

	Query Match	10.2%;	Score	713.4;	DB 1;	Length	1518;
	Best Local Similarity	67.7%;	Pred.	No. 2.4e-185;			
	Matches 1034;	Conservative	0;	Mismatches	481;	Indels	12;
	Gaps	2;					
Qy	3680	AAAAATGAACGCACCTATTGTGATTTGGCCAGCGCTTTGGCGCCTTGGCGCTTCGCCGATTC	3739				
Dd	2	AAACCATTGGAAAAAACCGCTTGTAATGGCGAGCGCTTTGGTGGCCTTGGCGCTTCGCCGATTC	61				
Qy	3740	GCCTGCAAGCGCGGGCATACCAACCACTTTACTTCGAGCAGCGCGACAACCCGGGCGGAC	3799				
Dd	62	GCCTGCAGCGCGCAGGGATCCCAACCGTACTGCTGGAGCAGCGGGACAAGCGCGGCGGTC	121				
Qy	3800	CGCGCTATGTTTTGAGGACAGTGGCTTTTACCTTCGATGCGGACCCACGGTGATCACCG	3859				
Dd	122	GGGCGCTACGCTGCGCATGAC CAGGGCTTTTACCTTTGACGCGGGCCGACGGTGATCACCG	181				
Qy	3860	ATCCCAGCGGCATCGAAGAGTTGTTCACGCTGGCAGGAAAATCGCTCAGCGAATTACGTCG	3919				
Dd	182	ATCCTACCGCGCTTGAGCGCGTGTTCACCTGGCGCGCAGCGCATGGAGGATTAAGTCGA	241				
Qy	3920	AGCTGATGCCGGTAACGCCCTTCTATCGCTGTGCTGGGAAGATGGCAAACAGCTTGATT	3979				
Dd	242	GGCTGCTCCGGTAAACCCCTTCTACCGACTCTGCTGGAGTCTCGGGGAAGACCCCTCGACT	301				
Qy	3980	ACGACATAATCAGCCGCTGCTGGAGCAGCATGCCAGTTCAAATCCGCAAGATGTAG	4039				
Dd	302	ATGCTAACGACAGCTTCGAGCTTGAGGCGCGAGTTTACCCAGTTCAACCCCCCGACGCTCG	361				



RESULT 9  
 US-08-096-623A-7  
 ; Sequence 7, Application US/08096623A  
 ; Patent No. 5684238  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ausich, Rodney L.  
 ; APPLICANT: Brinkhaus, Friedhelm L.  
 ; APPLICANT: Mukharji, Indrani  
 ; APPLICANT: Proffitt, John H.  
 ; APPLICANT: Yarger, James G.  
 ; APPLICANT: Yen, Huel-Che B.  
 ; TITLE OF INVENTION: Biosynthesis of Zeaxanthin and  
 ; TITLE OF INVENTION: Glycosylated Zeaxanthin in Genetically Engineered Hosts  
 ; NUMBER OF SEQUENCES: 104  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Welsh & Katz, Ltd.  
 ; STREET: 120 S. Riverside Plaza, 22nd Floor  
 ; CITY: Chicago  
 ; STATE: IL  
 ; COUNTRY: USA  
 ; ZIP: 60606  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION NUMBER: US/08/096,623A  
 ; APPLICATION NUMBER: US/08/096,623A  
 ; FILING DATE: 22-JUL-1993  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/805,061  
 ; FILING DATE: 09-DEC-1991  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/662,921  
 ; FILING DATE: 28-FEB-1991  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/562,674  
 ; FILING DATE: 03-AUG-1990  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/525,551  
 ; FILING DATE: 18-MAY-1990  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/487,613  
 ; FILING DATE: 02-MAR-1990  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Gamson, Edward P.  
 ; REGISTRATION NUMBER: 29,381  
 ; REFERENCE/DOCKET NUMBER: AMO-006.1  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (312) 655-1500  
 ; TELEFAX: (312) 655-1501  
 ; INFORMATION FOR SEQ ID NO: 7:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1518 base pairs

[illegible]



QY 4580 GTAAAGCGCATGAGCACTCGCTGTTTGTACTCTATTTTGGCCCTGATCAGCGCGCATGAC 4639  
Db 902 GCAAGAGCATGAGCACTCGCTGTTTGTACTCTATTTTGGCCCTGATCAGCGCGCATGAC 961  
QY 4640 AGCTCGCGCACCACACCGCTGTTTGGCCCGGTTATCGTCTAGTTGATCGATGAGATT 4699  
Db 962 AGCTGGCGCACCATACCATCTGTTTGGTCCCGCTACCGGAGCTGATCGAGAGATCT 1021  
QY 4700 TCAACAGCAGCAGCTGGCAGACGATTTTCTACTTCTGACCGCGCTCGCAGCGG 4759  
Db 1022 TTACCGGCGCAGCGCTGGCGGATGACTTCTCGCTCTACCTGACCTCGCCCTCGTGACG 1081  
QY 4760 ATCCGCTCGTGCACCGCGCGCTGCGGAGCTTTTATGTTAGCGCGGCTGCGCATC 4819  
Db 1082 ATCCCTCGTGCAGCGCTGCGCGCTGCGGAGCTTCTACGTTGCTGGCGCGGCTGCGCATC 1141  
QY 4820 TCGGCAACCGCTGACATGCACTGCAACAGAGGACCGCGCTTGGCGGATCGAATTTTG 4879  
Db 1142 TTGGCAACCGCGCTGCACTGCGCGCAGAGGGCGGAGCTGCGGACCGCATTTTG 1201  
QY 4880 CTTATCTGAGCAGCACTACATGCGGGATTAAGTCAGCAATTAAGTCAGCAGCAATGT 4939  
Db 1202 ACTACCTTGAAGAGCGCTATATGCGCGCTGCGTAGCCAGCTGGTAGCCAGCGATCT 1261  
QY 4940 TTACGCGCTTGTATTTTCGGCAGCAGCTGATGCCCATCAGCGCTCGGCTGTTTCTGCTGG 4999  
Db 1262 TTACCGCGCAGACTTTCACACACGCTTGGATCGGATCTTGGGATCGCTTTTTCATCGAGC 1321  
QY 5000 AGCCGATTTTGACGCAAGCGCTGTTTCCGCGCATTAACCGCGATGCGGATATCAGCA 5059  
Db 1322 CGCTTCGTTGACCAAGCGCTTGTTCGCGCAACCGCAGCAGCATTC-----NA 1372  
QY 5060 ATCTCTATCTGTTGGTGGCGGTACGATTCAGGCGCGCGCTGCGCGCGTGTATCGGTT 5119  
Db 1373 ACTCTACTGTTGGCGCAGGTACTACCTTGGCGCGGCAATCTCGCGGTAGTGGGCG 1432  
QY 5120 CGGCCAAGCCACCGCAGCTGATGCTGAGGATCGCGCGCAATGAATGACAGGCTTT 5179  
Db 1433 TCGCGAAAGCAGCGCGCTGATGATGAGGATC---TGCAATGAGCGCAACCGCGCT 1489  
QY 5180 ACTTGAGCAAGTAACGCAACCATGCG 5206  
Db 1490 GTTTGACGCGCAGCAGCAGCATGCG 1516

RESULT 10

US-08-095-726-9  
; Sequence 9, Application US/08095726  
; Patent No. 5530188  
; GENERAL INFORMATION:  
; APPLICANT: Ausich, Rodney L  
; APPLICANT: Brinkhaus, Friedrich L  
; APPLICANT: Mukharji, Indrani  
; APPLICANT: Proffitt, John H  
; APPLICANT: Yarger, James G  
; APPLICANT: Yen, Huel-Che B  
; TITLE OF INVENTION: Beta-Carotene Biosynthesis in  
; TITLE OF INVENTION: Genetically Engineered Hosts  
; NUMBER OF SEQUENCES: 79  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Amoco Corp., Patents and Licensing Dept  
; STREET: 200 E Randolph St  
; CITY: Chicago  
; STATE: IL  
; COUNTRY: USA  
; ZIP: 60680-0703  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.24  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/095,726

; FILING DATE: 21-JUL-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/785,566  
; FILING DATE: 30-OCT-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Galloway, No. 5530188val B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 3128567180  
; TELEFAX: 3128564972  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1522 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-095-726-9  
Query Match 10.1%; Score 708.4; DB 1; Length 1522;  
Best Local Similarity 67.6%; Pred. No. 5.6e-184;  
Matches 1029; Conservative 0; Mismatches 481; Indels 12; Gaps 2;  
QY 3685 ATGAAACGCACTTATGATTGGCGCAGCTTTGGCGCTGCGCTGCGGATTCGCGCTG 3744  
Db 11 ATGGAATAAACCGTTGATTGGCGCAGCTTTGGTGGCTTGGCTGCGGATTCGCGCTG 70  
QY 3745 CAAGCGGGCGGCATACCAACCACTTACTTCGAGCAGCGCGCAAAACCGGGCGGACGCGC 3804  
Db 71 CAGGCGGCGAGGATCCCAACCGTACTGCTGGAGCAGCGGACAGCCCGCGGTGGGCG 130  
QY 3805 TATGTTTGGAGCAGTGGCTTTTACCTTCGATGCGGACCGACCGTGTATCCAGATCCC 3864  
Db 131 TACGCTCGCATGACAGCGCTTTTACCTTCGCGCGGCGCGCGTGTATCCAGATCCC 190  
QY 3865 AGCGCCATCGAGAGCTTGTTCAGCTGCGCAGGAAATCGCTCAGCGATTAACGTCGAGCTG 3924  
Db 191 ACCTGCTGAGCGCTGTTTACCTTGGCGCGCAGCGCATGAGGATTAACGTCAGGCTG 250  
QY 3925 ATGCGCGTAAACCGCTTCTATCGCTGCTGCGGAAGATGGCAACAGCTTGAATACGAC 3984  
Db 251 CTGCGCGTAAACCGCTTCTATCGCTGCTGCGGAAGATGGCAACAGCTTGAATACGCT 310  
QY 3985 AATAATACGCGCTGCTGGAGCAGCAGATGCGCAATCGCAAGATGAGAGGCG 4044  
Db 311 AACGACAGCTTCAGCTTGGCGCGCAGATTAACCGCTTCAACCGCGCGCGCTGAGGCG 370  
QY 4045 TATCGCTCAATTTTGGCTTATTCAGTGAAGTATTTAGAGAGGTTATCTGAACTCGGC 4104  
Db 371 TACCGCGCTTCTGCTGCTTACTTCCAGCGGCTATTCAGAGGAGGATTTTGGCGCTCGGC 430  
QY 4105 ACCTGCGCGCTTCTGCGAGCTGCTGACATGCTGCGCGCTGCGCGCGCAGTTGGGACGCTG 4164  
Db 431 ACCTGCGCGCTTCTGCTGCTTCTGCGCAGATGCTGCGCGCGCGCGCGCTGCTTAACTC 490  
QY 4165 CAAGCATGCGCAGCGCTGCTACAGCAGTGGTGGCAATTTTATTCAGACGATCATCTGCGT 4224  
Db 491 CAGGCGTGGCAGAGGCTTACCGAGTGGTTCGCGCTTTTATTCAGGATGAGCATCTGCGG 550  
QY 4225 CAGGCGTTCCTTCCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4284  
Db 551 CAGGCGTTCCTTCCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 610  
QY 4285 TATACCTTAATTCATGCGCTGCGAGCTGAAATGGGCGCTGCTGCTGCTGCTGCTGCTG 4344  
Db 611 TACACCTGATCCAGCGCTTGGAGCGGAGTGGGGGCTGCTGCTGCTGAGGCGGCGCAC 670  
QY 4345 GCGCGCTGCTGCGAGGATGCGCGAGCTGTTTCAGAGGACTTGGGCGGCGAGCTGTTTACTG 4404  
Db 671 GCGGCGCTGCTGAGCGGATGCTGAGGCTGTTTACCGATCTGGGCGGCGGAGATCGAATC 730  
QY 4405 AATGCGGAGTGAAGCGCTGGAACACCGAGCTGGAACACCGAGCTTACGCGCTTCACTG 4464



Qy	4285	TATACCTTAATTCAATGCGCTGAGAGCGTGAATAGGGGCGCTGTGGTTTCGCGCGCGCGCAC	4344
Db	611	TACACCCCTGATCCACGCGCCCTTGAGCGGGAGTGGGGGCTCTGGTTCCCTGAGGGCGGCA	670
Qy	4345	GGGCGCTGTGTGACGGGCATGGCGCACTGTTTCGAGGACTTGGGCGCGAGCTGTACTG	4404
Db	671	GGGGCGCTGGTGAACGGCATGGTGAAGCTGTTTACCGATCTGGGCGGGGAGATCGAACTC	730
Qy	4405	AATGCCGAAGTGAGCCAGCTGGAAACAGCGGGCAATCGCATTAAGCGGCGTTCAGTTAGAG	4464
Db	731	AACGCCCGGTTCGAAGAGCTGGTGGTGGCCGATAAACCCGCTAAGCCAGGTCGCGCTCGG	790
Qy	4465	GGGGGACGACGCTTCGATGCGCGCGCTGTGGCTTCGAATCGGACGCTGTGTGCAATACCTAC	4524
Db	791	GATGGTCCGATCTTTTGACACGACGCGCTAGCCCTCGAAACGCTGACGCTGTGTGAACACCTAT	850
Qy	4525	GACAAACTGCTTCGCCACCATTCGCTGGCAATGAACGTCGGCATCGCTGAAGCGCTAAG	4584
Db	851	AAAAGCTGCTCGGCACCATACCGGTGGGCGAGAACGGGCGGACGGCTGGAGCGCAAG	910
Qy	4585	CGCATGAGCAACTCGCTGTTTGTACTCTATTTTGGCTGTAATCAGCGCGCATGAACAGCTC	4644
Db	911	AGCATGAGCAACTCGCTGTTTGCTCTACTTCGGCTGAACCAAGCTCAATCCAGCTG	970
Qy	4645	GCGGACACACGCTCTGTTTGGCCCGGTTATCGTGAGTTGATCGAATGAGATTTTCAAC	4704
Db	971	GCGCACCATACCATCTGTTTGGTCCCGCTACCGGGAGCTGATCGACGAGATCTTTACC	1030
Qy	4705	AGCAGCAGCTGGCAGACGATTTTCACTTTACTCTGACGCGGCTTCGACGAGATCCG	4764
Db	1031	GGCAGCGCTGGCGGATGACTTCTCGCTCTACTGCACTGCGGCTTCGTCGCGATCCC	1090
Qy	4765	TCGCTGGCACCGCCGCTCGGGCAGCTTTATGTGTAGCGCGGTGCGGCATCTCGGC	4824
Db	1091	TCGCTCGCGCTCCCGCTGGCCAGCTTCTAGTGTGCGCCCGGTGCGCATCTTGGC	1150
Qy	4825	ACCGTGACATCGACTGGGAAAGGAAAGACCGGCTTTCGCGATCGAATTTTGTCTTAT	4884
Db	1151	AACCGCGCTGGACTGGGCGCAGAGGGGCGGAAGCTGCGCAGCCGATCTTTGACTAC	1210
Qy	4885	CTGAGCAGCACTACATGCGGGATTAAGTACGCAATTAGTGACACACAGAAATGTTACG	4944
Db	1211	CTTGAGAGCGCTATATGCCCAGGCTCGGTAGCCAGCTGTGTGACCCAGCGGATCTTTACC	1270
Qy	4945	CCGTTTGATTTTCGCGACAGCTGATGCGCATACAGCGCTCGCGGTTTTCGCTGGAGCCG	5004
Db	1271	CGGAGACTTCAGACAGCTTGATCGCGATCTTGGGATCGCTTTTCATCGAGCGGCT	1330
Qy	5005	ATTTTGACGAAAGCGCTGGTTTCGCGCCCGATAACGCGGATCGGATATCAGCAATCTC	5064
Db	1331	TCGTTGACCAAGGCTTGTTCGCGCGAAAACGCGACACGACATTC-----AAACCTC	1381
Qy	5065	TATCTGTGGTGCCGTACGCATCCAGGCGCGGCGTTCGCGCGGTGATCGGTTGGGC	5124
Db	1382	TACTGTGTGCCGAGGATCTCAACCTTGGCGCGGACATTCCTGGCGTAGTGGGCGCTCGCC	1441
Qy	5125	AAGGCCACCGCAGGCTGATGCTGGAGGATCGGCCGAATGAATCGACAGCCTTTACTTG	5184
Db	1442	GAAGACACCGCAGCCTGATGATTGAGATC---TGCNATGAGCCAAACCGCGCTGCTTG	1498
Qy	5185	AGCAAGTAAACGCAAAACCATGGC	5206
Db	1499	ACCACGCCAGTCGACCATGGC	1520

**RESULT 12**

US-08-096-623A-9

; Sequence 9, Application US/08096623A  
; Patent No. 5684238

: GENERAL INFORMATION:

APPLICANT: Ausich, Rodney L.

APPLICANT: Brinkhaus, Friedhelm L.

APPLICANT: Mukharji, Indrani

APPLICANT: Proffitt, John H.  
 APPLICANT: Yarger, James G.  
 APPLICANT: Yen, Hwei-Che B.  
 TITLE OF INVENTION: Biosynthesis of Zeaxanthin and  
 TITLE OF INVENTION: Glycosylated Zeaxanthin in Genetically Engineered Hosts  
 NUMBER OF SEQUENCES: 104  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Welsh & Katz, Ltd.  
 STREET: 120 S. Riverside Plaza, 22nd Floor  
 CITY: Chicago  
 STATE: IL  
 COUNTRY: USA  
 ZIP: 60606  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/096,623A  
 FILING DATE: 22-JUL-1993  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/805,061  
 FILING DATE: 09-DEC-1991  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/662,921  
 FILING DATE: 28-FEB-1991  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/562,674  
 FILING DATE: 03-AUG-1990  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/525,551  
 FILING DATE: 18-MAY-1990  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/487,613  
 FILING DATE: 02-MAR-1990  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Gamson, Edward P.  
 REGISTRATION NUMBER: 29,381  
 REFERENCE/DOCKET NUMBER: AMO-006.1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (312) 655-1500  
 TELEFAX: (312) 655-1501  
 INFORMATION FOR SEQ ID NO: 9:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1522 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 POSITION IN GENOME:  
 MAP POSITION: -10 to 1512  
 UNITS: bp  
 US-08-096-623A-9

Query Match	10.1%;	Score 708.4;	DB 1;	Length 1522;
Best Local Similarity	67.6%;	Pred. No. 5.6e-184;		
Matches 1029; Conservative	0;	Mismatches 481;	Indels 12;	Gaps 2;

QY	3685	ATGAAACGCACTTATGTGATTGGCGCAGAGCTTTGGCGGGCTGGCGTGGCGATTCCGCTG	3744
Db	11	ATGGAAAAACCGTTGTGATTGGCGCAGGCTTTGGTGGCTGGCGATTCCGCTG	70
QY	3745	CAGCGCGGGGATACCAACACCTTACTCGACGAGGGGACAAAACGGGCGGACGGGCC	3804
Db	71	CAGCGCGCAGGGATCCCAACCGTACTGTGGAGCAGCGGACAAAGCCCGCGGTGGGCC	130
QY	3805	TATGTGTTTCAGACACAGTGACTTTTACCTTCGATGCCGACCCACGCTGATCCCGATCCC	3864
Db	131	TACGTCTGGGATGACACAGGGCTTTACCTTTGACGCGCGGGCCGACGCTGATCACCGATCCT	190
QY	3865	AGCGCCATCGAAGATGTGTTTCAGCTGGCAGGAAAAATCGCTCAGCGCATTACGTCGAGCTG	3924

Db	191	ACCGCGCTTGAGGCGCTGTTACCCCTGGCCGGCAGGCGCATGAGGAGTATCGTCAGGCTG	250
Qy	3925	ATGCGGTAAACGCCCTTCTATCGCTGTGCTGGGAAGATGGCAACAGCTTGATTAACGAC	3984
Db	251	CTGCGGTAAACCCCTTCTACCGACTCTGCTGGGAGTCCGGGAAAGACCTCGACTATGCT	310
Qy	3985	AATAATAGCGCGTCTCGAGCAGAGATCGCAGCTTCAATCCGCAAGATGTAGAAGGC	4044
Db	311	AACGACAGCTTCGAGCTTGAGCGCGAGATTACCCAGTTTCAACCCCGCGAGCTCGAGGCG	370
Qy	4045	TATCGTCAATTTCTGCGCTATTACGTCGAGTATTTAGAGAGGTTTCTCAAACTCGGC	4104
Db	371	TACCGCGCGTCTCTGCGCTTCTCCAGCGGTATTTCCAGAGGAGATATTTGGCGCTCGGC	430
Qy	4105	ACGCTGCCGCTTCTCGACGCTGCTGACATGCTGCGCGTTCGCGCAGATTTGGAGCTGTG	4164
Db	431	AGCGTGCGGTTCTCTCTCTTTTCGCGACATGCTGCGCGCGGCGCGAGCTGCTTAAAGCTC	490
Qy	4165	CAAGCATGGCGCAGCGTCTACAGCATGCTGGCGAAATTTATTCAGGACGATCATCTGCGT	4224
Db	491	CAGCGGTGGCAGAGCGTCTACAGCTCGGTTTCGCGCTTTTATTTAGAGATGAGCATCTGCGG	550
Qy	4225	CAGGCGTCTTCTTCCACTCATCTGCTGGTGGCGGTAAATCTTTTGCACCGTCAFCGATC	4284
Db	551	CAGGCTTCTCGTTCCACTCCCTGCTGCTGAGTAGCGGCAACCCCTTCAACACCTCGTCCATC	610
Qy	4285	TATACCTTAATTCATGCTGAGCGTGAATGGGGGTGTGTTTCCGGCGGCGGACCC	4344
Db	611	TACACCTGATCCAGCCCTTGAGCGGAGTGGGGGTCTGGTTCCCTTGAGGCGGCGACC	670
Qy	4345	GGCGCTGCTGAGGCGATGCGGACTGTTTCGAGGACTTGGCGGCGAGCTGTACTTG	4404
Db	671	GGGCGCTGCTGAACCGCATGCTGAAGCTTTTACCGATCTGGCGGGGAGATCGAATC	730
Qy	4405	AATGCCGAATGAGCGAGCTGAAACACGCGGAATTCGATTTAGCGCGCTTCAGTTAGAG	4464
Db	731	AACGCCCGGCTGAAGAGCTGTTGTTGGCGGATTAACCGCTTAAGCAGGCTCGCGCTCGG	790
Qy	4465	GGCGGACGAGCTTCATGCGCGCTGTGGCTCAATGCCGAGTGGTGTGATACCTAC	4524
Db	791	GATGGTGGATCTTTTACACCGAGCGCTGAGCTCGAAGCTGACGCTGTGTAACCTAT	850
Qy	4525	GACAACTGCTCGCCACCATCGCTGGCAATGAACGTCGACATCGCTGAAGCGTAAAG	4584
Db	851	AAAAGCTGCTCGCACCATACCGTGGTGGGCGAGAGCGGCGGCGAGCTGGAGCGAAG	910
Qy	4585	CGCATGAGCAACTCGCTGTTTGTACTATTTTGGCTGAAATCAGCGCGCATGAACGCTC	4644
Db	911	AGCATGAGCAACTCGCTGTTTGTGCTCTACTTCGGCTGAACACGCTCATTTCCAGCTG	970
Qy	4645	GCCACACACCTGCTGTTTGGCCCGGTTATCGTGAGTTGATCGATGAGATTTTCAAC	4704
Db	971	GGCACCATTACCATCTGTTTGTGCTCCCGCTACGGGAGCTGATCAAGAGATCTTTACC	1030
Qy	4705	AGCAGCAGCTGGCAGACGATTTTCACTTTTACTGCAAGCGGCTTCGAGAGCGATCGG	4764
Db	1031	GGCAGCGCTGGCGGATGATCTTCTGCTCTACTGCACTCGCGCTTCGCTGACCGATCCG	1090
Qy	4765	TCGCTGCAAGCGCGCTGGCGAGCTTTTATGTTAGCGCGGCTGGCGCATCTCGGC	4824
Db	1091	TCGCTCGCGCTCCCGCTGGCGAGCTTCTAGTGTGCGCCCGGTTGCGCATCTTGGC	1150
Qy	4825	ACCCTGACATCGACTGGCAACAGGAGGACCGGCTTTCGCGATCGAATTTTGTCTTAT	4884
Db	1151	AACCGCGCTGAGTCTGGCGCAGAGGCGCGAAGCTGCGCAGCATCTTTGACTAC	1210
Qy	4885	CTGAGCAGCATACATCGCGGATTAAGTACGATGAGCAATTAAGTACACACAGATTTTACG	4944
Db	1211	CTTGAAGAGCGCTATATGCGCGGCTGCTGAGCAGCTGTGTGAGCCAGCGGATCTTTACC	1270
Qy	4945	CCGTTTGTATTTTCGCGACACGCTGCATGCGCCATCACCGGCTCGGCGTTTTCGCTGGAGCCG	5004

Db	1271	CGCGAGACTTCACGACACACGCTTGGATCGGATCTTTGGGATCGCTTTTCATCGAGCGCCT	1330
Qy	5005	ATTTTCAGCAAGCGCTGTTTCGCGCGATACCGCGATGCCGATATCAGCAATCTC	5064
Db	1331	TCGTTGACCAAGGCTTGTTCGCGCAAAAGCGACACGATTC-----AAACCTC	1381
Qy	5065	TATCTGTGTGGTCCCGGTACGATTCAGCGCGGGCGTGCCTGGCGCTGATTCGGTTGGCC	5124
Db	1382	TACTGTGTGGCGGAGTACTACCTGCGCGGCAATCTCTGGCGTGTGGGCTCGCC	1441
Qy	5125	AAGCCACCGCGAGGCTGATGCTGGAGGATCGCGCGAATGATCGACAGCCTTTTACTTG	5184
Db	1442	GAAAGCACCGCGCTGATGATTTGAGGATC-----TGCAATGAGCCAAACCGCCCTGCTTG	1498
Qy	5185	AGCAAGTAACGCAAAACCATGCC	5206
Db	1499	ACACGCCACGTCACCATGCC	1520

RESULT 13  
US-08-660-645A-6  
; Sequence 6, Application US/08660645A  
; Patent No. 6087152  
; GENERAL INFORMATION:  
; APPLICANT: Hohmann, Hans-Peter  
; APPLICANT: Pasamontes, Luis  
; APPLICANT: Tessier, Michel  
; APPLICANT: van Loon, Adolphus  
; TITLE OF INVENTION: FERMENTATIVE CAROTENOID PRODUCTION  
; NUMBER OF SEQUENCES: 47  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hoffmann-La Roche Inc.  
; STREET: 340 Kingsland Street  
; CITY: Nutley  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 07110  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/660,645A  
; FILING DATE: 07-JUN-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: EP 9510888.9  
; FILING DATE: 09-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pokras, Bruce A.  
; REGISTRATION NUMBER: 32,748  
; REFERENCE/DOCKET NUMBER: RAN 6002/170  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (201) 235-5801  
; TELEFAX: (201) 235-2363  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1482 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; US-08-660-645A-6  
Query Match 8.8%; Score 619.4; DB 3; Length 1482;  
Best Local Similarity 63.7%; Pred. No. 1.5e-159;  
Matches 941; Conservative 0; Mismatches 536; Indels 0; Gaps 0;  
Qy 3685 ATCAACCGCATTGATTCGCGCAGGCTTTGGCGGCTTCGCGCTGCGGATTCGCTG 3744  
1 ATGAGTTCCGCATCGTCATCGCGAGGTTTCGCGGCTTCGCTTGCATCGGCTG 60  
Db



QY	3685	ATGAACGACCTATGATGAGGCGAGGCTTTGGCGGCTTGGCGCTGGCGATTCGGCTG	3744	QY	4765	TCGCTGGACCGCCCGGCTGCGGAGCTTTATGTATGTATGCTGCGCGGTCGCGCATCTCGGC	4824
Db	1	ATGAGTTCCGCCATCGTCAATCGGCGAGTTTCGGCGGCTTGGCGCTTGGCGATTCGGCTG	60	Db	1081	AGCTGGCCCGCGAGGGGATGTCACGATTAGTCTTGGCGCGTTCCGCGATCTGGGC	1140
QY	3745	CAAGCGGCGGAGATACCAACACCTTACTCAGACAGCGCGCAACAACCGGGGAGCGGCC	3804	QY	4825	ACCGTGAATCGATGGCAAGGACCGCGCTTGGCGGATCGAATTTTCTTAT	4884
Db	61	CAATCGGCGGATCGCGACCAACCATCGTCAGAGCCCGCGCAACCGCGGGCGCGGCC	120	Db	1141	CGCGCGATGTGATTTGGGAAGCGGCGCTATGCCGAGCGCATCTTCGAGGAA	1200
QY	3805	TATGTGTTGAGGACAGTGGCTTTACCTTCGATGCGGAGCCACCGGTGATCACCGATCCC	3864	QY	4885	CTGGAGCAGCATACATGCGGGATTTACGTACGAATTAGTGACACACAGAAATGTTTACG	4944
Db	121	TATGTCTGGAACGATCAGGCGACATCTCTTCGATGAGCGCGGACCGTCTGACCGACCC	180	Db	1201	CTGGAGCGCGCGCATCCCGACCTGCGCAAGCACTGACCGTCAAGCGCATCTTCAGC	1260
QY	3865	AGCGCCATCGAAGAGTTGTCAGCTGGCAGGAAATCGCTCAGCGATTAGTCGAGCTG	3924	QY	4945	CCGTTTGAATTTTCGCACACGCTGCAATCCGATCAGCGTTCGGCTTTCGCTGGAGCGG	5004
Db	181	GACAGCTCGAGAGCTGTGGGCGCTCAGCGGCAACCGATGAGCGGTGACGCTG	240	Db	1261	CCGCGGATTTTCAGCACCGAACTGTGCGGCCATCAGCGAGCGCTTCTCGGTTCGAGCGG	1320
QY	3925	ATGCGGTAACCGCTTCTATCGCTGTCTGCGAAGATGGCAACAGCTTGTATTCAGAC	3984	QY	5005	ATTTTACCGAAGCGCTGTTCCGCGCGATACCGGATGCGGATATCAGCAATCTC	5064
Db	241	CTGCGGCTCTCGCCCTTCTACCGGCTGACATGGCGGAGCGCGCAGCTTCGATACGTC	300	Db	1321	ATCTGACGCAATCCCGCTGGTTCCCGCCGATTAACCGGACCGCGCATCCGAACTTC	1380
QY	3985	AATAATCAGCGCTGCTGAGCAGAGATCGCCACGTTCAATCCGCAAGATGTAGAAGGC	4044	QY	5065	TATCTGTGTGGTTCGCGGTACGCAATCCAGGCGGGCGTTCGCGGCGTATCGGTTCCGCC	5124
Db	301	NACGACGACGAGCTGATCGCCAGCTCGCTCTTCAATCCCGCGGATGTTCGATGCG	360	Db	1381	TACATCGTGGGCGGCGACGCAATCCGCTGCGGCAATCCCGGCTGTCGTTGGCAGCGCC	1440
QY	4045	TATCGTCAATTTCTGCTTATTCAGTGAATTTAGAGGTTATCTGAACTCGGC	4104	QY	5125	AAGCCACCGCCAGGCTGATGCTGAGGATCGCGCG	5161
Db	361	TATCGCGCTTCACGATTTACGCGAGGAGGTCTATCGCGAGGGGTATCTGAAGCTGGG	420	Db	1441	AAGCCACCGCGCAGGTCATGCTGTCGACCTGGCGG	1477
QY	4105	ACGGTCCGTTCTGAGGCTGCTGACATGCTGCGGCTCGCGCGGAGTTGGAGCTCTG	4164	RESULT 15			
Db	421	ACCAAGCGCTTCTGAGCTGGCGAGATGCTGAACGCGCGCGCGCTGATGCGCTG	480	US-09-546-969-6			
QY	4165	CAAGCATGGCGAGCTCTACAGCATGGTGGGAAATTTATTCAGGACATCATCTGGT	4224	; Sequence 6, Application US/09546969			
Db	481	CAGGCAATACGCTCGCTCCACAGCATGGTGGCGCTTTCATCCAGGACCGCATCTCGGG	540	; Patent No. 6207409			
QY	4225	CAGCGTTTCTTCCACTCATGCTGGTGGGCGTAACTCTTTGCAACGTCATCGATC	4284	; GENERAL INFORMATION:			
Db	541	CAGCGCTTCTGTTTCACAGCTGCTGGTGGGCGGAAACCGTTCGACCGCTGATC	600	; APPLICANT: Hohmann, Hans-Peter			
QY	4285	TATACCTTAATTCATGCGCTGAGCGGTGAATGGGCGGTGTGGTTCCGCGCGCGCAC	4344	; APPLICANT: Pasamontes, Luis			
Db	601	TATGCGCTGATCCATGCGTGGAAACGCGCGCGGCTCTGGTTCCGCAAGCGCGCAC	660	; APPLICANT: Tessier, Michel			
QY	4345	GGCGGCTGTGAGGCGATGCGGACTGTTGAGGACTTGGCGCGGAGCTGTACTG	4404	; APPLICANT: van Loon, Adolphus			
Db	661	AACGAGCTGTGCGGCGATGGTGGCTTTCGAGGCTTTCGCGGCGACGCTGCTGCTG	720	; TITLE OF INVENTION: FERMENTATIVE CAROTENOID PRODUCTION			
QY	4405	AATCCGAGTGAGCCAGCTGGAAACGAGCGGCAATCGCAATTAGCGGCTTCAAGTTAGAG	4464	; NUMBER OF SEQUENCES: 47			
Db	721	AATGCGCGCTCAGCGGATCGACACGAGGCGATCGCGCCAGCGGCTCAAGCTGCTG	780	; CORRESPONDENCE ADDRESS:			
QY	4465	GGCGGACGAGCTTCGATGCGCGCTGTCGCTTCAATCCGCGGCTGGTGCATACCTAC	4524	; ADDRESSEE: Hoffmann-La Roche Inc.			
Db	781	GACGGGCGGAGTTGCGCGGAGATACGTTGGCCAGCAACCGGCGGATGATGACAGCTAT	840	; STREET: 340 Kingsland Street			
QY	4525	GACAACTGCTTCGCCACATTCGCTGCAATGAACTGCGCATCGCTGAAGCGGTAAG	4584	; CITY: Nutley			
Db	841	CGGACCTGTGGGCGATACCGCGCGGCGGCGACCAAGCGCGGATCTTGAACCGGAG	900	; STATE: NJ			
QY	4585	CGCATGAGCAACTCGCTTTTGTACTATATTTTGGCTTGAATCAGCGGATGAACAGCTC	4644	; COUNTRY: USA			
Db	901	CGTGTGATGTGCTGCTGCTGATTTTCGGCTGTCTCAAGCGCCCGGAGAACCTG	960	; ZIP: 07110			
QY	4645	GGGACACACGCTGTTTGGCGCGGTTATCGTGATTTGATCGATGAGATTTTCAAC	4704	; COMPUTER READABLE FORM:			
Db	961	GCCCCACACAGCTCATCTTGGCGCGGCTTACAGGGCTGTGTGACGAGATCTTCAAC	1020	; MEDIUM TYPE: Floppy disk			
QY	4705	AGCAGCGAGCTGGCAGAGATTTTCACTTTACTTCAAGCGCTTCAGCGGATCCG	4764	; COMPUTER: IBM PC compatible			
Db	1021	GGGCCACGCTTCCGCGACGATTTCTCGATGTATCTGATTCGCTTGGCGGATCCG	1080	; OPERATING SYSTEM: PC-DOS/MS-DOS			
				; SOFTWARE: Patentin Release #1.0, Version #1.30			
				; CURRENT APPLICATION DATA:			
				; APPLICATION NUMBER: US/09/546,969			
				; FILING DATE:			
				; CLASSIFICATION:			
				; PRIOR APPLICATION DATA:			
				; APPLICATION NUMBER: 08/660,645			
				; FILING DATE:			
				; ATTORNEY/AGENT INFORMATION:			
				; NAME: Pokras, Bruce A.			
				; REGISTRATION NUMBER: 32,748			
				; REFERENCE/DOCKET NUMBER: RAN 6002/170			
				; TELECOMMUNICATION INFORMATION:			
				; TELEPHONE: (201) 235-5801			
				; TELEFAX: (201) 235-2363			
				; INFORMATION FOR SEQ ID NO: 6:			
				; SEQUENCE CHARACTERISTICS:			
				; LENGTH: 1482 base pairs			
				; TYPE: nucleic acid			
				; STRANDEDNESS: double			
				; TOPOLOGY: linear			
				; MOLECULE TYPE: DNA (genomic)			



US-09-546-969-6

Query Match 8.8%; Score 619.4; DB 3; Length 1482;  
Best Local Similarity 63.7%; Pred. No. 1.5e-159;  
Matches 941; Conservative 0; Mismatches 536; Indels 0; Gaps 0;

```
QY 3685 ATGMAACGACCTATATGATTCGCGCAGAGCTTTGGCGGCTCGCGCTGGCGATTCGCCCTG 3744
Db 1 ATGAGTTCGCCATCGTCATCGCGCAGAGTTTCGGCGGGCTTCGGCTTCCCATTCGCCCTG 60

QY 3745 CAAGCGCGGGGATACCAACCACTTACTCGAGCAGCGCAAAACCGGGCGGAGCGCGC 3804
Db 61 CAATCGCGCGCATCGCACCAACCATCTCGAGGCGCGGACAGCGCGCGCGCGC 120

QY 3805 TATGTGTTTGGAGACAGTGGCTTTACTTCGATGCGCGGACCCACCGGTGATCAACGATCCC 3864
Db 121 TATGTGGAACGATCAGGCGCACAGCTTTCGATGACAGGCGCGACAGGTCTGTCGACCGACCC 180

QY 3865 AGCGCCATCGAAGAGTTGTCAGCTGGCGAGGAATCTCTCAGCGATTAACGTCGAGCTG 3924
Db 181 GACAGCTCGGAGAGCTGTGGGCCCTCAGGGCGCAACCGATGGAGCGTGACGCTG 240

QY 3925 ATGCCGCTAACGCCCTTCTATCGCTGTGTGGGAAGATGCAAAACAGCTTGATTAACGAC 3984
Db 241 CTGCGGCTCTCGCCCTTCTACCGGCTGACATGGCGGACGGCGCGAGCTTCGAACTAGTG 300

QY 3985 AATAATCAGCGCTGCTGAGCAGCAGATGCGCAAGTTCAATCCGCAAGATGTAGAAGC 4044
Db 301 AACGACGACGAGCTGATCGGCCAGGTCTGCTCAATCCCGCCGATCTCGATGGC 360

QY 4045 TATGTCGAATTTCTTCCCTTATTCAGTGAATTTAGAGAGGTTATCTGAACTCGGC 4104
Db 361 TATGCGCGCTTCACGATTAACGCGAGAGGTTATTCGCGAGGGGTATCTGAAGCTGGGG 420

QY 4105 AGCGTCCGCTTCTGAGAGTGGCTGACATGCTGCGCTCGCGCGAGTTGGAGAGCTGTG 4164
Db 421 ACCACGCGCTTCTGAGCTGGCCAGATGCTGAACGCGCGCGCGGCTGTATGCGCTG 480

QY 4165 CAAGCATGCGCAGCGCTTACAGCATGTTGGGGAATTTATTCAGGACGATCATCTGGT 4224
Db 481 CAGGCTACCGCTCGGTCCACAGCATGTTGGGCGCTTCATCCAGGACCGCATCTGCGG 540

QY 4225 CAGGCTTTTCTTCCACTCATGCTGTGGGGTAACTCTTTGCAACGTCATCGATC 4284
Db 541 CAGGCTTCTGCTTCACACGCTGTGCTGGCGGGAACCCGCTTTCGACCACTCGATC 600

QY 4285 TATACCTTAATTCATGCGCTGAGCGTGAATGGGCGTGTGTTTCCGCGCGCGCAC 4344
Db 601 TATGCGCTGATCCATGCGCTGGNAACGCGCGCGCGGCTGTGTTCCGCAAGGCGCGCAC 560

QY 4345 GCGCGCTGCTGGCAGGCGATGCGCGCATGTTTCGAGGACTTGGGCGGCGAGCTGTACTG 4404
Db 661 AACCACTGCTGGCGGATGTTGCGCTTTCGAGCGTCTTGGCGGACGCTGTGCTG 720

QY 4405 AATGCCGAAGTAGCCAGCTGGAAACAGCGGCAATCGCATTAGCGCGCTTCAGTTAGAG 4464
Db 721 AATGCCCGGCTACGCGGATGCAACCCGAGGCGGATCGCGCACCGGCGCTCAGCTGTG 780

QY 4465 GCGGACGACGCTTCGATGCGCGCTGTGGCTCCAAATGCGGACGTTGGTGCATACCTAC 4524
Db 781 GAGCGGCGGAGTTGTGGCGCGGATACGTTGGCCAGCAACGGGACGTTGATGCACAGCTAT 840

QY 4525 GACAACTGCTTGGCCACCATCGCTGGCAATGAACGTTGCGACATCGCTGAAGCGTAAG 4584
Db 841 CGGACCTGCTGGGCGCATACCGCGCGGCGGACCAAGGCGCGGATCTCGAACCGGCAG 900

QY 4585 CGCATGAGCACTCGCTGTTTGTACTTATTTTGGCTGAATCAGCGCGCATGAACAGCTC 4644
Db 901 CGCTGTGATGCTGCTGTGTTGCTGCTGCTTTCGAGGCTGTCCAAGCGCGCGGAGACCTG 960

QY 4645 GCGCACACACGCTGTTTGGCGCGGTTATCGTGAAGTGTATCGATCAGATTTTCAAC 4704
Db 961 GCGCACACACGCTGATCTTGGCGCGGCTTCAAGGGGCTGTGAACGAGATCTTCAAC 1020
```

Search completed: January 25, 2005, 18:28:48

Job time : 509 secs

```
QY 4705 AGCAGCCAGCTGGCAGACGATTTTTCATCTTACCTGACGCGCCCTGACAGCAGCATCCG 4764
Db 1021 GGGCCACGCTTCCCGACGATTTCTCGATGTATCTGCAATCGCTTCCGTCGACCGATCCC 1080

QY 4765 TCGCTGGCAGCCGCGCGCTGCGCAGCTTTTATGTTAGCCCGGTGCGCATCTCGGC 4824
Db 1081 AGCCTGGCCCCCGAGGGGATGTCACGCAATACGTCCTTGGCGCCGTTCCGATCTGGGG 1140

QY 4825 ACCGCTGACATCGACTGGCAACAGAAAGGACCGCTTTCGCGGATCGAAATTTTGTCTAT 4884
Db 1141 CGGCGCGATGTGATTTGGGAAGCCGAGGCCCGGCTATGCGAGCGCATCTTCGAGGAA 1200

QY 4885 CTGGAGCAGCATACTACATGCGCGGATTAAGTCAAGCAATTAAGTACACACAGAAATGTTTACG 4944
Db 1201 CTGGAGCGCGCGCATCCCGACCTGGCAAGACCTGACCGTCAGCCGCGCATCTTCAGC 1260

QY 4945 CGGTTTGTATTTTCGCGACACGCTGCAATGCCCATCAGGCTCGGCGTTTTCGCTGGAGCG 5004
Db 1261 CCGCGGATTTTCAGCACCGCACTGCGGCCCATCAGGCGAGCGCTTCTCGGTCGAGCG 1320

QY 5005 ATTTTGAACGAAAGCGCTTTCGCGCGCATAAACCGGATGCGGATTCAGCAATCTC 5064
Db 1321 ATCCTGACGCAATCCGCTTGGTTCCGCGCGCATAAACCGGACCGCGGATCCCGAATTC 1380

QY 5065 TATCTGTTGGTGGTCCGCTACGATCCAGCGCGCGGCTGCCCGCGGTGATTCGCTTCGCGC 5124
Db 1381 TACATCGTGGGGCGGCGACGATCCGGTGGCGCATCCCGGTGCTGCTGTCGTCGAGCGC 1440

QY 5125 AAGGCCACGCGCAGCTGATGCTGGAGGATCCGCGCG 5161
Db 1441 AAGGCCACGCGCAGCTGATGCTGCGGACCTGGGCGG 1477
```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (C) 1993 - 2005 Compugen Ltd.

QM nucleic - nucleic search, using sw model

Run on: January 25, 2005, 10:29:07 ; Search time 3350 Seconds  
(without alignments)  
12004.619 Million cell updates/sec

Title: US-10-808-979-18

Perfect score: 6999

Sequence: 1 gcgatggcaatggaaatgt.....gacgttttagacatgaata 6999

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.\*

1: /cgn2\_6/ptoddata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptoddata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptoddata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptoddata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptoddata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptoddata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptoddata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptoddata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptoddata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptoddata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptoddata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptoddata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptoddata/1/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptoddata/1/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptoddata/1/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptoddata/1/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptoddata/1/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptoddata/1/pubpna/US10\_NEW\_PUB.seq.\*  
19: /cgn2\_6/ptoddata/1/pubpna/US11\_NEW\_PUB.seq.\*  
20: /cgn2\_6/ptoddata/1/pubpna/US60\_NEW\_PUB.seq.\*  
21: /cgn2\_6/ptoddata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6999	100.0	6999	18	US-10-808-979-18
2	4326.2	61.8	8675	18	US-10-804-677-18
3	2619.6	37.4	8609	18	US-10-735-019-28
4	2619.6	37.4	8609	18	US-10-734-778-40
5	2619.6	37.4	8609	18	US-10-734-936-37
6	2619.6	37.4	8609	18	US-10-735-442-64
7	2138.6	30.6	8814	18	US-10-808-807-18
8	2135.2	30.5	12753	16	US-10-041-018-19
9	1618.8	23.1	7494	17	US-10-474-536-45
10	1618.8	23.1	8547	17	US-10-474-536-46
11	1482	21.2	1482	17	US-10-808-979-7
12	1278	18.3	1278	18	US-10-808-979-3

13	1167	16.7	1167	18	US-10-808-979-5	Sequence 5, Appli
14	1085.2	15.5	1482	18	US-10-804-677-7	Sequence 7, Appli
15	1035.4	14.8	8625	10	US-09-920-923-1	Sequence 1, Appli
16	1035.4	14.8	8625	16	US-10-695-980-1	Sequence 1, Appli
17	1035.4	14.8	11233	10	US-09-920-923-27	Sequence 27, Appli
18	1035.4	14.8	11233	16	US-10-695-980-27	Sequence 27, Appli
19	930	13.3	930	18	US-10-808-979-9	Sequence 9, Appli
20	906	12.9	906	18	US-10-808-979-1	Sequence 1, Appli
21	875.4	12.5	1278	18	US-10-804-677-3	Sequence 3, Appli
22	861.4	12.3	1479	18	US-10-808-807-7	Sequence 7, Appli
23	808.6	11.6	1479	10	US-09-941-947A-31	Sequence 31, Appli
24	808.6	11.6	1479	15	US-10-218-118-7	Sequence 7, Appli
25	808.6	11.6	1479	16	US-10-363-567-31	Sequence 31, Appli
26	808.6	11.6	1479	16	US-10-466-656-7	Sequence 7, Appli
27	808.6	11.6	1479	18	US-10-735-019-7	Sequence 7, Appli
28	808.6	11.6	1479	18	US-10-734-778-7	Sequence 7, Appli
29	808.6	11.6	1479	18	US-10-735-442-7	Sequence 7, Appli
30	808.6	11.6	1479	18	US-10-848-307-7	Sequence 7, Appli
31	695	9.9	1167	18	US-10-804-677-5	Sequence 5, Appli
32	692.6	9.9	1521	14	US-10-166-037-5	Sequence 5, Appli
33	630.8	9.0	930	18	US-10-804-677-9	Sequence 9, Appli
34	629.2	9.0	906	18	US-10-804-677-1	Sequence 1, Appli
35	619.4	8.8	1482	9	US-09-547-267-6	Sequence 6, Appli
36	535.6	7.7	1473	15	US-10-369-493-23921	Sequence 23921, A
37	534	7.6	534	18	US-10-808-979-11	Sequence 11, Appli
38	510.6	7.3	1518	15	US-10-369-493-41517	Sequence 41517, A
39	480.8	6.9	1494	15	US-10-369-493-34266	Sequence 34266, A
40	449.2	6.4	1296	10	US-09-941-947A-27	Sequence 27, Appli
41	449.2	6.4	1296	15	US-10-218-118-3	Sequence 3, Appli
42	449.2	6.4	1296	16	US-10-363-567-27	Sequence 27, Appli
43	449.2	6.4	1296	16	US-10-466-656-1	Sequence 1, Appli
44	449.2	6.4	1296	18	US-10-735-019-3	Sequence 3, Appli
45	449.2	6.4	1296	18	US-10-734-778-3	Sequence 3, Appli

#### ALIGNMENTS

RESULT 1  
US-10-808-979-18  
; Sequence 18, Application US/10808979  
; Publication No. US20040268439A1  
; GENERAL INFORMATION:  
; APPLICANT: E.I. duPont de Nemours and Company, Inc.  
; APPLICANT: Cheng, Qiong  
; APPLICANT: Tao, Luan  
; TITLE OF INVENTION: GENES ENCODING CAROTENOID COMPOUNDS  
; FILE REFERENCE: CL2360 US NA  
; CURRENT APPLICATION NUMBER: US/10/808,979  
; CURRENT FILING DATE: 2004-03-25  
; PRIOR APPLICATION NUMBER: US 60/471,904  
; PRIOR FILING DATE: 2003-05-20  
; PRIOR APPLICATION NUMBER: US 60/527,083  
; PRIOR FILING DATE: 2003-12-03  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 18  
; LENGTH: 6999  
; TYPE: DNA  
; ORGANISM: Enterobacteriaceae strain DC260  
US-10-808-979-18

Query Match	100.0%	Score 6999;	DB 18;	Length 6999;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 6999;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
QY	1	GGCATGGCAATGGAAATGTTCCGTGAGCTTTTCGTCTTAACTCGGGCATCAGCGGTGAA	60	
Db	1	GGCATGGCAATGGAAATGTTCCGTGAGCTTTTCGTCTTAACTCGGGCATCAGCGGTGAA	60	
QY	61	TAATCAGCACGCGAGGTTTGGTATTCGTCATTTTATGTCATTAGCGGTATTTATTAC	120	
Db	61	TAATCAGCACGCGAGGTTTGGTATTCGTCATTTTATGTCATTAGCGGTATTTATTAC	120	

Qy	121	CACAGCCTAACGAGTATTATATGTTACGGGTTGCTGTACTTATTTCCCTAATAAGA	180	1201	TTTAATCAACAGCTAGCGATATTCAACTAGAGCGCGCTCAGCCGCTGGGCCACTTTGCGG	1260
Db	121	CACAGCCTAACGAGTATTATGTTACGGGTTGCTGTACTTATTTCCCTAATAAGA	180	1201	TTTAATCAACAGCTAGCGATATTCAACTAGAGCGCGCTCAGCCGCTGGGCCACTTTGCGG	1260
Qy	181	TCACGATAGCAATTATTAACATATTTACCTGGTGCATGAATACGACACCTACAAAGT	240	1261	TGATCGCGCGCGCTCTACAGCACTTTCAACGCGTTGAGGGGTTAGCACAACCGCTGC	1320
Db	181	TCACGATAGCAATTATTAACATATTTACCTGGTGCATGAATACGACACCTACAAAGT	240	1261	TGATCGCGCGCGCTCTACAGCACTTTCAACGCGTTGAGGGGTTAGCACAACCGCTGC	1320
Qy	241	CAAGTCCCTCGCTGGCGAACTCACCTTAGCGAGTCTACGGTTTAATCAAAAAAGCATAAAA	300	1321	TGCGCGCGCGCCATCGCATCACATTCATCCAGCAAGCCGATGCCGCACCTTTGCTTAGCG	1380
Db	241	CAAGTCCCTCGCTGGCGAACTCACCTTAGCGAGTCTACGGTTTAATCAAAAAAGCATAAAA	300	1321	TGCGCGCGCGCCATCGCATCACATTCATCCAGCAAGCCGATGCCGCACCTTTGCTTAGCG	1380
Qy	301	TTTACCAACCAATGATAGCCATATGACCAACCAATGTCACACACAGCAGCATCAGACA	360	1381	ACGAACGATCGATTTGTTGCGCTCGGCAACAGACGATCTGCTGCCGGTTGCTGCGCG	1440
Db	301	TTTACCAACCAATGATAGCCATATGACCAACCAATGTCACACACAGCAGCATCAGACA	360	1381	ACGAACGATCGATTTGTTGCGCTCGGCAACAGACGATCTGCTGCCGGTTGCTGCGCG	1440
Qy	361	AGCAACTCTCTCAGCTGCAGCAAAATTTACAGCGCATCTTGAACATTTACTGCTGCC	420	1441	CCGTGTTGATCGGCTGGCCTCGCGGGCGGCTGCTGCTGTTTTCGCTGATCGACGATC	1500
Db	361	AGCAACTCTCTCAGCTGCAGCAAAATTTACAGCGCATCTTGAACATTTACTGCTGCC	420	1441	CCGTGTTGATCGGCTGGCCTCGCGGGCGGCTGCTGCTGTTTTCGCTGATCGACGATC	1500
Qy	421	GGACAGCAAGGATCGCTGCGTGCAGATGCGTGCGGAAACGCTGGCGCAGGGCAAA	480	1501	TCGCTCTCTGCACCGATATGCTGTGCCGCAACTGTGCTGCGTACTGAAAGCATTTGAACA	1560
Db	421	GGACAGCAAGGATCGCTGCGTGCAGATGCGTGCGGAAACGCTGGCGCAGGGCAAA	480	1501	TCGCTCTCTGCACCGATATGCTGTGCCGCAACTGTGCTGCGTACTGAAAGCATTTGAACA	1560
Qy	481	CGTATTCGTCCTTATTACTGCTGCGCAGCGCGATGATGGTTGCGAGCTGACGCAA	540	1561	TCGATGGCGTGAATCGCCGACGAAATGGAAGGGGGCGGATTTGGTCTGTAAGCGCTGC	1620
Db	481	CGTATTCGTCCTTATTACTGCTGCGCAGCGCGATGATGGTTGCGAGCTGACGCAA	540	1561	TCGATGGCGTGAATCGCCGACGAAATGGAAGGGGGCGGATTTGGTCTGTAAGCGCTGC	1620
Qy	541	AATGGCGTCTCGATCTCGCTGTGAGTGGAAATGGTGACCGCGCATTCGTGATTCG	600	1621	ATCTGCCGTTTGTTCGCTGGCCTGCGCTTTCGCGTCAATCGTGAAGCGGGATTCGCG	1680
Db	541	AATGGCGTCTCGATCTCGCTGTGAGTGGAAATGGTGACCGCGCATTCGTGATTCG	600	1621	ATCTGCCGTTTGTTCGCTGGCCTGCGCTTTCGCGTCAATCGTGAAGCGGGATTCGCG	1680
Qy	601	GATGACATTCCTCGATGATACCGCGAGTGCCTGCTGCTGCGCATCGCGCATTCGATTCG	660	1681	TTGCGGTGATGCCCTTCCGTTTTCGACAGGATGCAAGAGCGCTGAAACGTTTTCAGGCCA	1740
Db	601	GATGACATTCCTCGATGATACCGCGAGTGCCTGCTGCTGCGCATCGCGCATTCGATTCG	660	1681	TTGCGGTGATGCCCTTCCGTTTTCGACAGGATGCAAGAGCGCTGAAACGTTTTCAGGCCA	1740
Qy	661	GAATTTGGTGAAACGTTGGGATTTCTCGCGCATCGCGCTGCTTAGCGCGCATTTGAA	720	1741	GCAGGATATCTATGATCGCATGCTGTGTCAGCGGAGCTGATCCTCAAAACACGCGC	1800
Db	661	GAATTTGGTGAAACGTTGGGATTTCTCGCGCATCGCGCTGCTTAGCGCGCATTTGAA	720	1741	GCAGGATATCTATGATCGCATGCTGTGTCAGCGGAGCTGATCCTCAAAACACGCGC	1800
Qy	721	GTGATGTCATTCGACCGGTTTTCGCTGCGCATATGAAATCTGAAGCGATTCGTAATC	780	1801	GGGCGTTTAAATTTGACAGGAGCGCGGATTTACATAGTGCCTGTGCGCGTGCACAAA	1860
Db	721	GTGATGTCATTCGACCGGTTTTCGCTGCGCATATGAAATCTGAAGCGATTCGTAATC	780	1801	GGGCGTTTAAATTTGACAGGAGCGCGGATTTACATAGTGCCTGTGCGCGTGCACAAA	1860
Qy	781	TCCGCTGCGCTGCGCATGAGGCTTAGTCAAGGCAATTCAGGATCTGACACGCGC	840	1861	TCAGCAGATGGTGC CGGCTTTGATTTTCCAGCTCAGCACTGCCGCTGCTATCACG	1920
Db	781	TCCGCTGCGCTGCGCATGAGGCTTAGTCAAGGCAATTCAGGATCTGACACGCGC	840	1861	TCAGCAGATGGTGC CGGCTTTGATTTTCCAGCTCAGCACTGCCGCTGCTATCACG	1920
Qy	841	ACGACAGCGCGAGCGCGGAGCGATCGCCATGACCAAGAACTGAAACACGAGCTGCTG	900	1921	CCGTGGGCGCACTCGCGCGCGCTTCTCTGCGCGCTCCATGCGCGCTTGGCCAGCGC	1980
Db	841	ACGACAGCGCGAGCGCGGAGCGATCGCCATGACCAAGAACTGAAACACGAGCTGCTG	900	1921	CCGTGGGCGCACTCGCGCGCGCTTCTCTGCGCGCTCCATGCGCGCTTGGCCAGCGC	1980
Qy	901	TTTCGCGCGAGCTGCAAAATGGCGGATTTGCGCTGACGCTTCACGAGGTGGCGCAA	960	1981	TGCGTACGCGCGGTTTATGCTGCTGCGGTACGTCGCAAGGCGCATCGCTTCGCGCTGT	2040
Db	901	TTTCGCGCGAGCTGCAAAATGGCGGATTTGCGCTGACGCTTCACGAGGTGGCGCAA	960	1981	TGCGTACGCGCGGTTTATGCTGCTGCGGTACGTCGCAAGGCGCATCGCTTCGCGCTGT	2040
Qy	961	AGACTTAGCTTCTTCGCGCGAGGATTTGGGCGAGCGTTTCACTGCTGCAAGCACTCGCC	1020	2041	TTTCGATCTGCGCAGCGGTGCCCGCAGCTGCGGCTATCGCTGATCGCCCATTTGTG	2100
Db	961	AGACTTAGCTTCTTCGCGCGAGGATTTGGGCGAGCGTTTCACTGCTGCAAGCACTCGCC	1020	2041	TTTCGATCTGCGCAGCGGTGCCCGCAGCTGCGGCTATCGCTGATCGCCCATTTGTG	2100
Qy	1021	GACGTTTGCACCAACACCGGTAAAGATGTGACCAAGGATCAGGCGCAATCCACGCTGGTA	1080	2101	GGGGATTAACCGCGCAACAGACGATCAGCTGGAGCTCGCTGGCGCGCTGGGTGACGG	2160
Db	1021	GACGTTTGCACCAACACCGGTAAAGATGTGACCAAGGATCAGGCGCAATCCACGCTGGTA	1080	2101	GGGGATTAACCGCGCAACAGACGATCAGCTGGAGCTCGCTGGCGCGCTGGGTGACGG	2160
Qy	1081	CAGATGCTCGTGTGAGCGCGCGGACCTGCGCTGCGCATCACCTGCGCAGCGAGAT	1140	2161	ATTTGCTGATCAGCGCGCGCTTACAGCGCGCTGTTTATCTACTCATGCGCGGT	2220
Db	1081	CAGATGCTCGTGTGAGCGCGCGGACCTGCGCTGCGCATCACCTGCGCAGCGAGAT	1140	2161	ATTTGCTGATCAGCGCGCGCTTACAGCGCGCTGTTTATCTACTCATGCGCGGT	2220
Qy	1141	GCACACCTTTCGCTGCGCTGCGCATCGCGGATCGCCATCGCCAAATATATGACGCGCTG	1200	2221	TAAACAGCGCGCTGGAAGCACTGGAATCGGTACCGCATGCTGCGCTGCGCATTTGCTT	2280
Db	1141	GCACACCTTTCGCTGCGCTGCGCATCGCGGATCGCCATCGCCAAATATATGACGCGCTG	1200	2221	TAAACAGCGCGCTGGAAGCACTGGAATCGGTACCGCATGCTGCGCTGCGCATTTGCTT	2280
				2281	TTGATCAGCGCGCTGGCGCGCATTTAGTGGCATGACGTTGCTGCGCGCGCATCAC	2340

[illegible]

Db	3361	GGCGCTAGCGGAGTGTGGTAGCAGCGCTGTGTGCCACCGATGCCCTCAGCGCTCAGCCAAACA	3422
Qy	3421	TATCGAACGCTTTTGCCCGTCAGCAGTGGCGCAACAGCGATTTTCCGTCGTCTAAACCG	3480
Db	3421	TATCGAACGCTTTTGCCCGTCAGCAGTGGCGCAACAGCGATTTTCCGTCGTCTAAACCG	3480
Qy	3481	CATGCTGTTTTTGCCCGGTAAGCGCAGCAGCGCTGGCGGTGATCAACGTTTTTATCCG	3540
Db	3481	CATGCTGTTTTTGCCCGGTAAGCGCAGCAGCGCTGGCGGTGATCAACGTTTTTATCCG	3540
Qy	3541	GCTCGATGCCGGGTTAATTAGCGCTTTTACGCCGGGCAACTCGCGCTCGCGCATAAAC	3600
Db	3541	GCTCGATGCCGGGTTAATTAGCGCTTTTACGCCGGGCAACTCGCGCTCGCGCATAAAC	3600
Qy	3601	GCAGATTCTCTGCGGCAAGCGCGCTGCCATCTCGTGAAGCGCTGCGCGCGCTGTGAA	3660
Db	3601	GCAGATTCTCTGCGGCAAGCGCGCGCTGCCATCTCGTGAAGCGCTGCGCGCGCTGTGAA	3660
Qy	3661	TTCTGTGCAACCAAGGGAAGAAAAATGAAACGCATTTATGTGATTGGCGCAGGCTTTGGC	3720
Db	3661	TTCTGTGCAACCAAGGGAAGAAAAATGAAACGCATTTATGTGATTGGCGCAGGCTTTGGC	3720
Qy	3721	GGCTGTGGCGCTGGCGATTCCGCTGCAAGCGCGGGGATACCAACGACCTTACTCGAGCAG	3780
Db	3721	GGCTGTGGCGCTGGCGATTCCGCTGCAAGCGCGGGGATACCAACGACCTTACTCGAGCAG	3780
Qy	3781	CGCGACAAACCGGGCGAGCGGCTATGTGTTTTGAGACAGTGGCTTACCTTCGATGCC	3840
Db	3781	CGCGACAAACCGGGCGAGCGGCTATGTGTTTTGAGACAGTGGCTTACCTTCGATGCC	3840
Qy	3841	GGACCAACGCTGATACCGATCCAGCGCCATCGAAGAGTGTTCACGCTGGCAGGAAAA	3900
Db	3841	GGACCAACGCTGATACCGATCCAGCGCCATCGAAGAGTGTTCACGCTGGCAGGAAAA	3900
Qy	3901	TCGCTCAGCGATTACGTGAGCTGATGCCGGTAACGGCCTTCTATCGCTGTGCTGGGAA	3960
Db	3901	TCGCTCAGCGATTACGTGAGCTGATGCCGGTAACGGCCTTCTATCGCTGTGCTGGGAA	3960
Qy	3961	GATGGGAAAACAGCTTCGATTACGACAATATACAGCGCTGTGAGCAGCAGATCGCCACG	4020
Db	3961	GATGGGAAAACAGCTTCGATTACGACAATATACAGCGCTGTGAGCAGCAGATCGCCACG	4020
Qy	4021	TTCAATCCGCAAGATGTAAGGCTATCGTCAATTTCTTGCCCTATTCACTGAAGTATTT	4080
Db	4021	TTCAATCCGCAAGATGTAAGGCTATCGTCAATTTCTTGCCCTATTCACTGAAGTATTT	4080
Qy	4081	AGAGAGGTTATCTGAAACTCGGCAAGTGGCGGTTCTGCAAGTGGTGATCTGTCGC	4140
Db	4081	AGAGAGGTTATCTGAAACTCGGCAAGTGGCGGTTCTGCAAGTGGTGATCTGTCGC	4140
Qy	4141	GTCCGCGCGAGTTGGGACGCTGCAAGCATGCGCAGCGCTCTACAGCATGTTGGCGAAA	4200
Db	4141	GTCCGCGCGAGTTGGGACGCTGCAAGCATGCGCAGCGCTCTACAGCATGTTGGCGAAA	4200
Qy	4201	TTTATTACGACGATCATCTGCGTCAGGCGTTTTCCCTTCCAATTTGCTGGTGGCGGT	4260
Db	4201	TTTATTACGACGATCATCTGCGTCAGGCGTTTTCCCTTCCAATTTGCTGGTGGCGGT	4260
Qy	4261	AATCCTTTTGCAACGTCATCGATCTATACTTAAATTCATGCGCTGGAGCGTGAATGGGC	4320
Db	4261	AATCCTTTTGCAACGTCATCGATCTATACTTAAATTCATGCGCTGGAGCGTGAATGGGC	4320
Qy	4321	GTCTGTTTTCCGCGCGGCGCACCGCGGCTGGTCGAGGCGATGCGGACGTTTCGAG	4380
Db	4321	GTCTGTTTTCCGCGCGGCGCACCGCGGCTGGTCGAGGCGATGCGGACGTTTCGAG	4380
Qy	4381	GACTTGGCGCGGAGCTGTTACTGAATGCCAAGTGAGCGAGCTGGAACACGCGGCAAT	4440
Db	4381	GACTTGGCGCGGAGCTGTTACTGAATGCCAAGTGAGCGAGCTGGAACACGCGGCAAT	4440
Qy	4441	CGAATTAGCGGCTTCAGTTAGAGGCGGACGACGCTTCGATGCGCGCTGTGGCTCC	4500
Db	4441	CGAATTAGCGGCTTCAGTTAGAGGCGGACGACGCTTCGATGCGCGCTGTGGCTCC	4500

QY 4501 AATGCCGAGCTGGTGCATACCTA CGA CAA AACTGCTTCGCCACCATCCGCTGGCAATGAAA 4560  
DB |||||  
QY 4501 AATGCCGAGCTGGTGCATACCTA CGA CAA AACTGCTTCGCCACCATCCGCTGGCAATGAAA 4560  
DB |||||  
QY 4561 CQTGCGACATCGCTCAAGCGTAAGCGCATGAGCAACTCGCTGTTGTACTCTATTTTGGC 4620  
DB |||||  
QY 4561 CQTGCGACATCGCTCAAGCGTAAGCGCATGAGCAACTCGCTGTTGTACTCTATTTTGGC 4620  
DB |||||  
QY 4621 CTGAATCAGCCGATGAACAGCTCCGCGCA CCA CACCGTCTGTTTGGCCCGGTTATCGT 4680  
DB |||||  
QY 4621 CTGAATCAGCCGATGAACAGCTCCGCGCA CCA CACCGTCTGTTTGGCCCGGTTATCGT 4680  
DB |||||  
QY 4681 GAGTTGATCGATGATATTTTCAACAGCAGCAGCTGGCAGACGATTTTCTACCTTACCTG 4740  
DB |||||  
QY 4681 GAGTTGATCGATGATATTTTCAACAGCAGCAGCTGGCAGACGATTTTCTACCTTACCTG 4740  
DB |||||  
QY 4741 CACGCGCCCTGCAGCAGCAGATCCGTCGCTGGCACCGCCGCTGGCGAGCTTTTATGTG 4800  
DB |||||  
QY 4741 CACGCGCCCTGCAGCAGCAGATCCGTCGCTGGCACCGCCGCTGGCGAGCTTTTATGTG 4800  
DB |||||  
QY 4801 TTAGCGCCGCTGCAGCAGCAGATCCGTCGCTGGCACCGCTGACATGCACTGGCAACAGGAGCGCGC 4860  
DB |||||  
QY 4801 TTAGCGCCGCTGCAGCAGCAGATCCGTCGCTGGCACCGCTGACATGCACTGGCAACAGGAGCGCGC 4860  
DB |||||  
QY 4861 TTGCGGATCGAATTTTTCGCTTATCTGGAGCAGCACTACATGCCGGGATTTAGCTCAGCAA 4920  
DB |||||  
QY 4861 TTGCGGATCGAATTTTTCGCTTATCTGGAGCAGCACTACATGCCGGGATTTAGCTCAGCAA 4920  
DB |||||  
QY 4921 TTAGTGACACACAGAAATTTTACGCGTTTGTATTTTGGCGACACGCTGATGCCCATCAC 4980  
DB |||||  
QY 4921 TTAGTGACACACAGAAATTTTACGCGTTTGTATTTTGGCGACACGCTGATGCCCATCAC 4980  
DB |||||  
QY 4981 GGCCTGGCGTTTTCGCTGAGCGGATTTTGA CGCAAAGCGCTGTTCCGCCCGCATAC 5040  
DB |||||  
QY 4981 GGCCTGGCGTTTTCGCTGAGCGGATTTTGA CGCAAAGCGCTGTTCCGCCCGCATAC 5040  
DB |||||  
QY 5041 CGCGATGCCGATATCAGCAATCTCTATCTGGTGGGTGCGGTCAGCATCCAGCGCGGGC 5100  
DB |||||  
QY 5041 CGCGATGCCGATATCAGCAATCTCTATCTGGTGGGTGCGGTCAGCATCCAGCGCGGGC 5100  
DB |||||  
QY 5101 GTGCCCGGCTGATCGGTTCCGCAAGGCCACCGCAGCGCTGATGCTGAGGATCGCGC 5160  
DB |||||  
QY 5101 GTGCCCGGCTGATCGGTTCCGCAAGGCCACCGCAGCGCTGATGCTGAGGATCGCGC 5160  
DB |||||  
QY 5161 GAATGAATCGACAGCCTTTACTTTGAGCAAGTAACGCAACCACTGGCGTGGCTCGAAGA 5220  
DB |||||  
QY 5161 GAATGAATCGACAGCCTTTACTTTGAGCAAGTAACGCAACCACTGGCGTGGCTCGAAGA 5220  
DB |||||  
QY 5221 GTTTCCGCAACCGCCGCAAGCTGTTTATGTCACCGACCGCGCGCAGCACGCTGATGCTGT 5280  
DB |||||  
QY 5221 GTTTCCGCAACCGCCGCAAGCTGTTTATGTCACCGACCGCGCGCAGCACGCTGATGCTGT 5280  
DB |||||  
QY 5281 ATCGGTGGTGTCTGCTACCTGGAGATGATGATGATGCGCAACCTGGCGGAGCGGCA 5340  
DB |||||  
QY 5281 ATCGGTGGTGTCTGCTACCTGGAGATGATGATGATGCGCAACCTGGCGGAGCGGCA 5340  
DB |||||  
QY 5341 CGCAGCATCGCTCGAAGACCGCAGCGCATGATGAGCATCTGCAAAATTTGAACCCCGC 5400  
DB |||||  
QY 5341 CGCAGCATCGCTCGAAGACCGCAGCGCATGATGAGCATCTGCAAAATTTGAACCCCGC 5400  
DB |||||  
QY 5401 GCGCCTACAGCGCGCGCACATGATGAA CCGCGCTTTAGGCGCTTCAGGAAAGTGGCGA 5460  
DB |||||  
QY 5461 TCATTACAGCTGCCGCAACACTGGCTTTGATCATCTGGAGGCTTCGCTATGATG 5520  
DB |||||  
QY 5461 TCATTACAGCTGCCGCAACACTGGCTTTGATCATCTGGAGGCTTCGCTATGATG 5520  
DB |||||  
QY 5521 CAGCAACGAACTTACGAGCTTCGATGACACGCTGCGTTACTGCTATCAGCTCGCG 5580  
DB |||||  
QY 5521 CAGCAACGAACTTACGAGCTTCGATGACACGCTGCGTTACTGCTATCAGCTCGCG 5580  
DB |||||

QY 5581 GCGTGTGCTGTTGATGATGGCGCGGCTAATGGCGCTGCGCGACGAGCGGTCTCGATC 5640  
DB |||||  
QY 5581 GCGTGTGCTGTTGATGATGGCGCGGCTAATGGCGCTGCGCGACGAGCGGTCTCGATC 5640  
DB |||||  
QY 5641 AGCGCTGCGATTTAGGACTGGCGTTCCAGCTCACTAACTTGGCGCGGCACTGTAGAAG 5700  
DB |||||  
QY 5641 AGCGCTGCGATTTAGGACTGGCGTTCCAGCTCACTAACTTGGCGCGGCACTGTAGAAG 5700  
DB |||||  
QY 5701 ATGCCGAAATGCTGCTGCTATCTGCCGCAATCTCTGGCTCGATCAGGCGGGATTTACGCG 5760  
DB |||||  
QY 5701 ATGCCGAAATGCTGCTGCTATCTGCCGCAATCTCTGGCTCGATCAGGCGGGATTTACGCG 5760  
DB |||||  
QY 5761 CCGATACGCTGACTGACCCGCAACATCGTACGAGCGCTCGCTCACTGCGAGCGGTTAG 5820  
DB |||||  
QY 5761 CCGATACGCTGACTGACCCGCAACATCGTACGAGCGCTCGCTCACTGCGAGCGGTTAG 5820  
DB |||||  
QY 5821 TGGCGAGGCGGAACCTATATCACTCGGCGGATCCGCTTTTACCGGTTTACCGCTGC 5880  
DB |||||  
QY 5821 TGGCGAGGCGGGAACCTATATCACTCGGCGGATCCGCTTTTACCGGTTTACCGCTGC 5880  
DB |||||  
QY 5881 GCTCGGCTGGGCGCATCGCTACGCGCTCGCGCGTATTCGCGAAATTTGGCGTCAAAGTTC 5940  
DB |||||  
QY 5881 GCTCGGCTGGGCGCATCGCTACGCGCTCGCGCGTATTCGCGAAATTTGGCGTCAAAGTTC 5940  
DB |||||  
QY 5941 AGCA CGCGGCTGTCACGCTGGGATTCACGCGAGCGCACAGTAAGAGTGGAACCTGG 6000  
DB |||||  
QY 5941 AGCA CGCGGCTGTCACGCTGGGATTCACGCGAGCGCACAGTAAGAGTGGAACCTGG 6000  
DB |||||  
QY 6001 CGCTGCTGTGAAAGGCGAGTTTGGCGATCACTCGCGTGTCTCTGCTCTGAAACCGC 6060  
DB |||||  
QY 6001 CGCTGCTGTGAAAGGCGAGTTTGGCGATCACTCGCGTGTCTCTGCTCTGAAACCGC 6060  
DB |||||  
QY 6061 GTCCGCGTGTCTGTGCGAGCGTCTCTGTTGATTTTACGTCGCTGACGCTGGCGCAGCGT 6120  
DB |||||  
QY 6061 GTCCGCGTGTCTGTGCGAGCGTCTCTGTTGATTTTACGTCGCTGACGCTGGCGCAGCGT 6120  
DB |||||  
QY 6121 GCGTTCAGCTTATTCAGCGGTGCGGTAGAGGAACCAACACAGCGAGCTTCACG 6180  
DB |||||  
QY 6121 GCGTTCAGCTTATTCAGCGGTGCGGTAGAGGAACCAACACAGCGAGCTTCACG 6180  
DB |||||  
QY 6181 CCGCGCACCGCATGATGCTGCGCGCATGTATAGCGCTTAAGCGCTTAAGATAGCTTTGCG 6240  
DB |||||  
QY 6181 CCGCGCACCGCATGATGCTGCGCGCATGTATAGCGCTTAAGCGCTTAAGATAGCTTTGCG 6240  
DB |||||  
QY 6241 CCGGATATAGCGGAACCGCGCATGATGACACAGGCGCATCGTACCACTGAGTAGAG 6300  
DB |||||  
QY 6241 CCGGATATAGCGGAACCGCGCATGATGACACAGGCGCATCGTACCACTGAGTAGAG 6300  
DB |||||  
QY 6301 CCGCGCTAGCTGCTCATTCGCGCACCAATCCACTGACGCGCGCACATGCTTGCACACC 6360  
DB |||||  
QY 6301 CCGCGCTAGCTGCTCATTCGCGCACCAATCCACTGACGCGCGCACATGCTTGCACACC 6360  
DB |||||  
QY 6361 GACATAAATCAGCACAATCGCCAGTACCGCAAAACCA CCGCATGAGATCGTTGAGCTC 6420  
DB |||||  
QY 6361 GACATAAATCAGCACAATCGCCAGTACCGCAAAACCA CCGCATGAGATCGTTGAGCTC 6420  
DB |||||  
QY 6421 AAACCTACCGCTGCTGCTCATGCTGCGACAGATGCGAGCGGCGGCGGCGGCGGCGG 6480  
DB |||||  
QY 6421 AAACCTACCGCTGCTGCTCATGCTGCGACAGATGCGAGCGGCGGCGGCGGCGGCGG 6480  
DB |||||  
QY 6481 GATGTATTTATGCGACAGCGCGCTTACGATTTTCCATCCACCGTTCGCAACAGAT 6540  
DB |||||  
QY 6481 GATGTATTTATGCGACAGCGCGCTTACGATTTTCCATCCACCGTTCGCAACAGAT 6540  
DB |||||  
QY 6541 AAGCACTTCCATAACCGAGAGATTTGCTGCTCATTTTGGAAAGGAGTACTAAAGG 6600  
DB |||||  
QY 6541 AAGCACTTCCATAACCGAGAGATTTGCTGCTCATTTTGGAAAGGAGTACTAAAGG 6600  
DB |||||  
QY 6601 TCGACGCGGATGAGTGAGGCGCAAGTTTACATGTTTAGAAATTTTAAAGTCCATAA 6660  
DB |||||  
QY 6601 TCGACGCGGATGAGTGAGGCGCAAGTTTACATGTTTAGAAATTTTAAAGTCCATAA 6660  
DB |||||  
QY 6661 CACGCTTATGAACGCTGCTATTCAGAAAGCGAGATTTTCA CACATACTCACCACTTATC 6720  
DB |||||





Db 2499 CGTTAAGTCGCGTTCGTTGCGGTGGCTGCGCGCTCCCGTCAATCGCGAGGGCCA 2558  
Qy 1675 TTCCGCTGCGGTGATGCCCTTCGTTTGGCAGAGATGACAAAGCGCTGAAGCGTTTTC 1734  
Db 2559 TTCCGCTGCGGTGATGCCCTTCGTTTGGCAGAGATGACAAAGCGCTGAGCGCTATC 2618  
Qy 1735 AGGCAGAGCGGATATCTATGATCGCATCGCTGCTGCTGCGGACGCTGATCCTCAAC 1794  
Db 2619 AGGCAGAGCGGATATCTATGATCGCATCGCTGCTGCTGCGGACGCTGATCCTCAAC 2678  
Qy 1795 AGCGCGGCGGTTTAAATTTAGCGAGCGCGCGGATTAATCAGTGCCTGTCGCGCTGG 1854  
Db 2679 ATGCGCGCGCTTCGCGCTGCCGAAAGCCATGGCTTATACATGATGCTGCTGCGCGCTGG 2738  
Qy 1855 CAAATACAGCAGATGTCGCGGCTTTGATTTTCCAGTCAAGCACTGCCCGCTGCT 1914  
Db 2739 CGCAATCAGTCAGCTGCTGCGCGCTTTGATTTTCCAGCGCAGCAACTGCCAGCTGCT 2798  
Qy 1915 ATCAGCGGTGGGCACTCCGCGCCCGGTTTCTCCTGCGCGCTCATGCGCCCTGGC 1974  
Db 2799 ATCAGAGCTGGGTCGCTGCGGACTCCAGTTGCTAGCGCGCGCTCGCGCACTCGCTGGC 2858  
Qy 1975 CAGCGCTGCGTCAAGCGGTGTTTATGCTCGCTGCGTACGTGCAAGCCATCGCTTCC 2034  
Db 2859 CAGCGCTGCGCAAGCGGTGTTTATGCTCGCTGCGTACGTGCAAGCCATCGCTTCC 2918  
Qy 2035 GCGTGTTCGTGATTCGGCGAGCGGTGCGGCACTGCGGCTATCGCTGATGATCGGCC 2094  
Db 2919 GCGTGTTCGTGATTCGGCGAGCGGTGCGGCACTGCGGCTATCGCTGATGATGATG 2978  
Qy 2095 ATTGTCGGGATTAACGCGCAACAGACGATCAGCTGAGCTGCTGCGCGCGGTGGG 2154  
Db 2979 ACTGTGCGGTTGACCGCGCAGCGGCACTCAGCTCAGCTGCGCGGTGCTGGTGG 3038  
Qy 2155 TGACGATTTGCGTATGATGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2214  
Db 3039 TGACGATTTGCGTATGATGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3098  
Qy 2215 CCGGTTTAAACGCGCGCTGGAAGCACTGGAATGCGGTACGCGGATGCTGCGCGTCCGA 2274  
Db 3099 CCGGTTTAAACGCGCGCTGGAAGCACTGGAATGCGGTACGCGGATGCTGCGCGTCCGA 3158  
Qy 2275 TTGCTTTTGAATCAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2334  
Db 3159 TCGCCTTCGATCAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3218  
Qy 2335 CATCAGCGTTTACGCGGTTCATCAACTGAGGAGCAGCATCTGCAACAGCTGCTGACGAG 2394  
Db 3219 CCTCAGCTTTTACGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3278  
Qy 2395 ATCGTTACGCGTACGAGTGTACGCGATTCAGCGGATTCAGCGGCGGCGGCGGCGGCGG 2454  
Db 3279 ATCGCTATGCTTGGCGATGTACGCGATTCAGCGGATTCAGCGGCGGCGGCGGCGGCGG 3338  
Qy 2455 AGCGTCCGCGGATCGTCAGAGCGGCTGTGCGAGCGAGCAAGTGTGCTGCGCGGAGG 2514  
Db 3339 CGCGCGGCGGCTGATTTGTGCGAGCGGCGGCTGTGTCAGCAGCAAAATCGTGTGCGGAGG 3398  
Qy 2515 CGACCTGATGCGGCACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2574  
Db 3399 CCACCTGATGCGGCACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3458  
Qy 2575 GATTGCGCTGCTGCGTCAATTTGACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2634  
Db 3459 GATTGCGCTGCTGCGTCAATTTGACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3518  
Qy 2635 TGCGCATCGCGGAGGCAATCATACCTGCTGCTTTCATCAGCGGATCTCAGCGCGGCA 2694  
Db 3519 GGCGCAGCGCGGCGGCAATCATACCTGCTGCTTTCATCAGCGGATCTCAGCGCGGCA 3578  
Qy 2695 ACTTCGCTGGTGCACCGCTGATTAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2754

Db 3579 GTTTCGCTGCTCGAGCGGCTGCTTTCGCGCGCTGCGCGCGGTTATCAGGTACGCTTCCC 3638  
Qy 2755 TGGCTGCGCGCAATCTGAGCGGGATTAATTTCCATCGCATCAGGCGATTTTCCCG 2814  
Db 3639 CACCTGCGTGGCGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3698  
Qy 2815 CCATCTTTACGCGCGATGGTGACATCTGTGGGAAACACAGCGGTACACAGGTAAA 2874  
Db 3699 GCATTACAGCAGGTGCTGCGTGGCGCTACGACCGAGCGCGGTCAGCGAGGTCTC 3758  
Qy 2875 ACCACGAGGTGACGCTGCGGATGGCGGTGAATCTGTCGCAAGTGGTGAATGATGG 2934  
Db 3759 ACCACGCGGTGAGACTGCGGATGGCGGATGTTACAGCGAGCGGCTGATGACGG 3818  
Qy 2935 TCGCGGCTCGAGCGGACGACATCTGACGCTGGGTTATCAGGTGTTTCTTCGACAAGA 2994  
Db 3819 ACGCGGCTGAGCGGACACCGCATCTGACGCTCGGCTATCAGGCATTTTTCGGTCAGGA 3878  
Qy 2995 GTGCGAGCTGGCGAGCGGCGAGCGCTGAGAGCGGATCCTGATGATGATGATGATGATGATG 3054  
Db 3879 GTGCGAACTGGCGCGCGCATGCGCTGCGAGCGCAATATGATGATGATGATGATGATGATG 3938  
Qy 3055 TCAGCAGCGGTTATCGTTTGTCTACAGCTGCGCTCAGCGCGGATCGGCTATTGAT 3114  
Db 3939 TCAGCAGCGGTTATCGTTTGTCTACAGCTGCGCTCAGCGCGGATCGGCTATTGAT 3998  
Qy 3115 TGAAGATACCCATTTACGTTTAAACGAGCGCGCTGGGAGAAACACCGCTCGTAGGACAT 3174  
Db 3999 TGAAGATACCCATTTACGTTTAAACGAGCGCGCTGGGAGAAACACCGCTCGTAGGACAT 4058  
Qy 3175 GCGGATGATGCGCAATCAGCAAGCTGAGCGCTGAGTACGCTGCTGCTGAGAGCAGCG 3234  
Db 4059 TACGAGTTATGCCCAACAGCGCGGCTGGAAATTTGCGGAGCTGCTGCGAGGAGCAGCG 4118  
Qy 3235 CATATTACCGATTACCTGAGCGGCAACATCGATTCGATTCGCGCAACAGCAGCGCGGCA 3294  
Db 4119 CTCGCTGCGGATCAOGCTCAGCGCGGATATCGATCAGTTCTTGGCAACAGCAGCAGCGCA 4178  
Qy 3295 AGCGTGAAGCGGCTGCGCGCGGCTGTTTCAAGCCACCGGTTACTTCTTCCGCTC 3354  
Db 4179 ACCGTGAGCGGCTGCGCGCGGCTGTTTCAAGCCACCGGTTACTTCTTCCGCTC 4238  
Qy 3355 CGCGTGGCTGAGCGGATTTGAGAGCGCTGTTGCCACCGGATCGCTCAGCGCTCAG 3414  
Db 4239 CGCGTGGCTGAGCGGATTTGAGAGCGCTGTTGCCACCGGATCGCTCAGCGCTCAG 4298  
Qy 3415 CCAACATATCGAAGCTTTTCCGCTGAGAGTGGCGGAGAACAGCGATTTTTCGCTGCT 3474  
Db 4299 CCACTGCTGAAATCCTTTGCCCGCTCAGCACTGGCGGAGAGCGCTTTTTCGCTGCT 4358  
Qy 3475 AAAACCGGATGCTGTTTGGCGGCTTAAAGCGAGCGCTGGCGGATGAGTGAACGTTT 3534  
Db 4359 AAATCGCATGCTGTTTGGCGGCTTAAAGCGAGCGCTGGCGGATGAGTGAACGTTT 4418  
Qy 3535 TTACCGGCTCGATGCGGCTTAAATTTAGCGGCTTTTACCGCGGCAATGCGGCTCAGCGA 3594  
Db 4419 TTACCGGCTTGAAGCGGATTTAGCGGCTTTTACCGCGGCAATGCGGCTCAGCGA 4478  
Qy 3595 TAAACCGGATGCTGCGGCAAGCGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 3651  
Db 4479 TAAAGCAGCATTTCTGTGCGCAACCGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 4538  
Qy 3652 GCTGTTGATTTCTGTCGAACAGGAGAAATGAAACGCACTTATGATGATGCGGCA 3711  
Db 4539 GATGATGATCTTCCGTTTACAGGAGAAATGAAACGCACTTATGATGATGCGGCA 4598  
Qy 3712 GCGTTTGGCGCTGCGGCTGCGGATTCGCTGCAAGCGGCGGATTAACACCGCTT 3771  
Db 4599 GCGTTTGGCGCTGCGGCTGCGGATTCGCTGCAAGCGGCGGCTGCGGCTGCGGCTGCGGCT 4658  
Qy 3772 CTCGAGCAGCGGCAACCGCGGCGGCTGCTGTTTGGAGCAGTGGCTTTTACC 3831  
Db 4659 CTGGAACAGCGGATTAAGCTTGGCGGCGGCTGCTGTTTGGAGCAGTGGCTTTTACC 4718



Db 6879 AAAAGCTGCGCTGCTGTTGAAAGGTGCGGGCTGCGGCTTACTTTCGGCGCTTGTCTATC 6938  
QY 6052 CTGAACCGCGCTCGGCTGTCTGTGGCAGCGTCTCTGTGTGATTTTACGTTCGTGTGACGCTG 6111  
Db 6939 CCGAGCGCGCTCTGTCGGCTCTGTGTGACGCTCGCGTGTGACACGACGCCATGCGCGCTG 6998  
QY 6112 GCGCAGCGGTGCTGTCAGCTTATTCAGCGGTGCGCGTAGAGAAACCAAAACGACACGCA 6171  
Db 6999 GCGCAGCGTGGCTGCGCTGCTGTAACCGTGGGCGTAAAGAAAGCCAAAGAAACGCA 7058  
QY 6172 GCGCTTACCGCGCGCACCGCATGATGATGCGGTGCGCCATGTATTAAGCGCTTAAGATA 6231  
Db 7059 GCGCTCCCGTCCCGCACCGCATGATGATGCGGTGCGCCATATACAAACCGCTTACAGATA 7118  
QY 6232 GCGCTTGGCGGGATATAGCGAAACCGCGAGGTGATGACACAGGCCATCGTGACCCAT 6291  
Db 7119 GCGCTTGGCGTGAATATAGCGAAACCGCGAGGTGATGAGACAGGCCGTATGACCCAT 7178  
QY 6292 GAAGTAGAGCGCGCTGCTGCTCATTTCCGCGCACCAATCCACTGACGGGCGCACATGCC 6351  
Db 7179 AAAATAGAGAGCGCCATAGTGTCTATGCGAGGCGCATCCACTGACGGCGCCAGCGCC 7238  
QY 6352 TTGCACACCGCATAAATCAGCACAAATGCGCAGTACCGCAACACACCGCATAAAGATC 6411  
Db 7239 GTTGACACCGCAAAATCAGGCCAATCGAACACAGCGCAACACACCGGCATACAGGTC 7298  
QY 6412 GTTGAGCTCAAACTTACCCTGTGCGGTTTCATGTTGCGACAGATGCCAGCCCATCCCA 6471  
Db 7299 GTTGAGCTCAAACTTACCCTGTGCGGTTTCATGTTGCGACAGATGCCAGCCCATCCCA 7358  
QY 6472 ACCGTGCATGATGATTTATGGACAGCGCGCTACGATTTCCATCACCACCAACGCGTTGC 6531  
Db 7359 GCCATGCAATGATTTATGGACAGCGCGGAGCATCTCCATCAGTATACCGGTAGC 7418  
QY 6532 CACAAAGATGACGCTTCCATTAACAGAGCATGTTGTCCTCATTTGTGTAAGAAAGGAG 6591  
Db 7419 CAGCAGGATAAGCGCATTCATTAACAGAGCATCATTTGTTCCATTTTGGCAGAGTGAAG 7478  
QY 6592 TACTAAGGTGAGCGGATGAGTGTGCGCAAG 6626  
Db 7479 TA-TAAAGGTGAGCGTGGATAGCGAAAGCGCAAG 7512

RESULT 3

US-10-735-019-28  
; Sequence 28, Application US/10735019  
; Publication No. US20040191863A1  
; GENERAL INFORMATION:  
; APPLICANT: E. I. duPont de Nemours and Company, Inc.  
; APPLICANT: Cheng, Qiong  
; APPLICANT: Suh, Wonchul  
; TITLE OF INVENTION: Mutations Affecting Plasmid Copy Number  
; FILE REFERENCE: CL2029 US NA  
; CURRENT APPLICATION NUMBER: US/10735, 019  
; CURRENT FILING DATE: 2003-12-12  
; PRIOR APPLICATION NUMBER: US 60/434973  
; PRIOR FILING DATE: 2002-12-20  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 28  
; LENGTH: 8609  
; TYPE: DNA  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Plasmid pPCB15  
US-10-735-019-28  
Query Match 37.4%; Score 2619.6; DB 18; Length 8609;  
Best Local Similarity 64.7%; Pred. No. 0;  
Matches 4015; Conservative 0; Mismatches 2169; Indels 24; Gaps 7;  
QY 373 CAGCTGCAGCAAAATTTACAGGCGCATCTTGACATTTACTGCTCGCGGACAGCAAGC 432

Db 1627 CAGTTGCTGGCTGATATCGATAGCCGCTTGTATCAGTTACTGCGCGTTCCAGGTCAGCGG 1686  
QY 433 GATCGCTGCGTGGCGGATGCGGTGCGGACCGTGGCGGAGGCAACGATTTTCGTCT 492  
Db 1687 GATTTGTGTGGGTGCGGATGCGGTGAAGGACGCTGGCACCGGGCAACGATTTTCGTCCG 1746  
QY 493 TTATTTACTGCTGCTGGCAGCGCGCATATGAGTTGCGAGCTGACGCAAAATGGCGTTCTC 552  
Db 1747 ATGCTGCTGTTATTAACAGCGCGCATCTTGGCTGTGGCATCAGTCACGGGGGATTA 1806  
QY 553 GATTCGCTGCTGTCAGTGGAAATGGTGCACCGCGCATGCTGATCTGGATGACATTTCCC 612  
Db 1807 GATTTAGCTGCGCGTGAATGGTGCATGCTGCTGCTGATTTCTGGATGATATGCC 1866  
QY 613 TCGATGATTAACCGCGCATGCGTGGTGGCTTACCGCTACCGTGCATCGCGAATTTGGTGAA 672  
Db 1867 TGCATGAGACGATGCGCGATGCGTGGGGGCGTCCCACCATTCACACGATGACGTTGAA 1926  
QY 673 AACGTGGCGATTTCTGCGCGCATGCGCTGTTAGCGCGCATTTGAAAGTGAATTTGCCATT 732  
Db 1927 CATGTGGCGATTTCTGGGGCGGTGCTTTACTCAGCAAAAGCGTTTGGGGTGAATTTGCCGAG 1986  
QY 733 GCACCGGTTTGGCTGCGCATACATATAATCTGAAAGGATTTGTAACCTCTCGCTGCGCTC 792  
Db 1987 GCTGAAGGTCTGACGCGCATAGCCAAACTCGCGCGGTGTCGAGCTGTCTCAGTGCAT 2046  
QY 793 GGCCTGACAGGCTTAGTSCAAGGGCAATTCAGAGATCTGCACGCGGCAACGACAGCGCGC 852  
Db 2047 GGCATGCAAGGTGTGTTGAGGGCGAGTTTAAAGACCTCTCGGAAGGCGATTAACCCCGC 2106  
QY 853 AGCCCGGAAGCGATGCGCATACCAACGAACTGAAACACGAGCTGTCTTTCGGCGCAAG 912  
Db 2107 AGCGCGGATGCTACTCTAAACCAATCAGTTTAAACACGACGCTGTTTGGCGCTCA 2166  
QY 913 CTGCAATTTGGCGGATTTGCGCTGACGCTTACCGAGGTGCGGCAAGACTTTAGCTTC 972  
Db 2167 AGCAAAATGGCGTCCATTTGCGGCCAACCGCTCTGCGAAGCGCGTGAGAACCTGCACTGT 2226  
QY 973 TTGCGCCAGAGATTTGGGCGAGCGCTTCAACTGCTCGACGACCTCGCGCGAGGTTGCAAA 1032  
Db 2227 TTCTGCTCGATCTCGGCGAGCGCTTTCAGTTGCTTGACGATCTTACCGATGGCATGACC 2286  
QY 1033 CACACCGTGAAGATGTGCACAGATCAGGGCAATCAGCTGCTGGTACAGATGCTCGT 1092  
Db 2287 GATACCGCAAAAGACATCAATCAGGATGAGTAAATCAACGCTGCTCAATTTATTTAGGC 2346  
QY 1093 GCTGACGGCGGCAACGCTGCGCTGCGATCACCTGCGCAGCGAGATGACACCTTGC 1152  
Db 2347 TCAGCGCGGTGCAAGAACGCTGCGACAGCATTTTGGCGCTTGGCAGTGAACACCTTTCC 2406  
QY 1153 TGCGCTGCCATCGCGCATCGCCACTCGCCAA---TATATGACGCGCTGTTTAAATCAA 1209  
Db 2407 GCGGATGCAAAACGCGCATTTCCACACCACTTTTATTCAGGCTGTTGACAAA 2466  
QY 1210 CAGCTAGCGATTTTCACTGAGCGCGGTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 1269  
Db 2467 AAATCTGCTGCGCTGAGTAAAGGATGTGCA-----TGAGCCATTTTGGCGTGA 2521  
QY 1270 GCGCGCTCTACAGCCACTTTTCAAGCGTTTGCAGCGGTTAGCACAACGCTGCTGGCGCG 1329  
Db 2522 GCGCCCTTTTTCAGCCATGTTTCGGCTCTGCAAAACCTTGTCTCAGGAATTTAGTGGCGCG 2581  
QY 1330 GCCATCGCATCATTTTCAAGCAAGCGGATGCCCGCATTTGCTTACGCGACGAAACGGA 1389  
Db 2582 GTCATCGTGTACGTTTTCAGCAACATGACTGCAAAAGCGCTGTTAAAGCGGACGAGATA 2641  
QY 1390 TCGATTTTGTGCGCTGCGGCAACAGAGCATCTTCCGCTGCGGCTGCGGCTGCGGCTGCGG 1449  
Db 2642 TCGGATTCAGACCGCTGCGCAACGATCTCTCCGCTTCTTATTCGCACTGCTGCTG 2701  
QY 1450 ATCGCTGCGCTCGCGCGCGCTGCTGCTGTTTCGCGTGTATCGACGATCTCGGCTCCT 1509

Db 2702 ACCTGGCGCGCACCACTCGGACCCCTCGATGTTAGCACTGATCAATGAATGGCAGTA 2761  
Qy 1510 GACCGATATGTCGCGCGCAACTGCTGTCGGGTACTGAAGCATGTAACATPCATGCGG 1569  
Db 2762 CCAGCGATATGCTTTGGCGGAACTGCCCGCGCTTTTCATGCGTTCAGATAGAGGCG 2821  
Qy 1570 TGATCGCGACGAATGGAAGCGCGCGGAGTGGTCTGCTGAAGCGTGCATCTGCCGT 1629  
Db 2822 TGATCGTTGATCAATGAGCGCGGAGGTGCAATGTCGCAAGCGTCGAGTCTGCCGT 2881  
Qy 1630 TTGTTTCGGTGGCTGCGCTTTCGGCTCAATCGTGAAGCGCGGATTCGGTTGCGGTGA 1689  
Db 2882 TTGTTTCGGTGGCTGCGCTGCGCTGCGCTCAACCGCGAACCAGGTTTGCTCTGGCGGTGA 2941  
Qy 1690 TGCCCTTCGGTTTGGACAGGATGAACAAAGCGCTGAAGCGTTTTCAGGCGCAGCAGGATA 1749  
Db 2942 TGCCCTTCGAGTACGCGCACGAGGATCGCGCTCGGGAACGCTATACCAACAGCGAAGAAA 3001  
Qy 1750 TCTATGATCGCATATGCTGCTACGCGGAGTGCATCTCAAAACAGCGCGGGGTTTA 1809  
Db 3002 TTATGACTGGCTGATGCGAGCTACGATCGGTGATCGCGCATCATGATGTCAGAAATGG 3061  
Qy 1810 ATTTGACGAGCGCGCGGATTTACATCAGTGCCTGTCGCGCTGCGCAACAAATCAGCCAGA 1869  
Db 3062 GTTTAGCCCCGGTGAAGAACTGTCATCTGTTTCTCCACTGCGCAAAATCAGCCAGT 3121  
Qy 1870 TGGTCCGGCTTTGATTTTCACTGACAGAACTGCCCGCTGCTATCAAGCGCTGGGCG 1929  
Db 3122 TGATCCCGMACTGGATTTTCCCGCAAGCGCTGCCAGCTGCTTTTCATGCGGTGGAC 3181  
Qy 1930 CACTTCGGG---CCCCGGTTTCTCTGCGCGCTCATGCGCCCTGGCCAGGCGTGCCTC 1986  
Db 3182 CGTTACGCGAACCCAGGGGAGCGCGGGGTCACTCAACTTCTTATTTTCCGTCGCCGACA 3241  
Qy 1987 AGCCGGTGGTTATGCTGCTGGGTGACGCTCAAGGCGCATCGCTTCGCGCTGTTTCTGC 2046  
Db 3242 AACCCGATATTTTGGCTGCTGGGCAACCTGACGGGACATGTTATGGCTGTTTCAGGA 3301  
Qy 2047 ATCTGGCGAGCGTGGCGCCAGCTGCGGCTATGCTGCTGATGCTCCCATGTTGGGGAT 2106  
Db 3302 CCATCGCCAAAGCCTGCGAAGAGTGGATGCGCAGTTACTGTTGGCACACTGTGGCGGCC 3361  
Qy 2107 TAAACGCGACAGAGCATCAGCTGAGCTGCTGGCGCGCTGGGTGACGATTTGCG 2166  
Db 3362 TCTAGCCACGAGGAGTGTAATGCGCCCGGGCGGGGACATTCAGGTTGTTGGATTTTG 3421  
Qy 2167 TCGATCAGCGCGACGCTTACAGCACGCGCAGCTGTTTATCACTCATGCGCGGTTTAAACA 2226  
Db 3422 CGGATCAATCCGACGACTTTTCAGAGGACAGTTGACATCAACATGTTGGGATGAATA 3481  
Qy 2227 GCGCGCTGGAAGCACTGGAATGCGGTACGCGATGCTGGCGTGGCGGATTTGTTGATC 2286  
Db 3482 CGGTACTGGACGCTATTGCTTCCCGCACACGCGTACTGGCGCTGCGCTGGCATTTGATC 3541  
Qy 2287 AGCCCGGCTGGCGCGGCTTATGATGGCATGAGTGGTGGTGGCGCGCGCATCACGCTTTA 2346  
Db 3542 AACCTGGGCTGGCATCAGCAATTTGTTTATCATGGCATCGGCAAGCGTGGCTCTCGGTTA 3601  
Qy 2347 GCGGTGTTTCACTCACTGGAGCAGCATCTGCAACAGCTGCTGACCGACGATCGTTACGCGC 2406  
Db 3602 CTACAGCCATCGCTGGCGCGCAGATTCGATCGCTGCTGACTAACACCGATTAACCCG 3661  
Qy 2407 TACGATGTACGATTTACGGCGCAGCTGACGCGCGCAGGCGGTTGCGCAGCGTGGCGCG 2466  
Db 3662 AGCGTATGACAAATTCAGGCGCATTTGCGTCTGGCAGGCGCACACGAGCGCGCGCG 3721  
Qy 2467 ACATCGTCGAGCAGCGCTGTCGACGACGACAGTGTGCTGGCGGAGGGGACCTAGGG 2526  
Db 3722 ATATTTGTTGAACAGGCGATGCG- GACCTGTCTAGCCAGTACTCAGTGGGCGGATTTATGCA 3780  
Qy 2527 CACGCAATACGATGATTTTGGTCTGGTCTGAGCTGGCGAATGCTGATTTGCGCTGCG 2586  
Db 3781 ACCGCACTATGATCTCAATCTGGTGGTGGCGGCTGTTGGCTAATGGCTTTATGCGGCTCG 3840

Qy 2587 TCTGGTCAATTTGAGCGCACAACTGAATGCTGTTGCTGAGAGCGGATGCGCATCCGCG 2646  
Db 3841 GCTTCAGCAACAGCATCCGGATATCGGATCTTGCTTATGAGGCGGCTCTGAGCGCG 3900  
Qy 2647 AGGCAATCATACCTGGTCTTTTCATCAGAGGATCTCAGCGCGCAACAACTTCGCTGGCT 2706  
Db 3901 AGGGAACCATACCTGGTCTTTTCAGGAAGAGATTTAAAGCTGAATCAGCATCGCTGGAT 3960  
Qy 2707 GCAACCGCTGATTTACCGTGGCTTGGTCAAGTTTATCAGGTGGGTTTTCCTCGCTCGCGCG 2766  
Db 3961 AGCGCGCTTGGTTCATCACTGGCCGACTACAGGTTCGTTTCCCGCAACGCGCTCG 4020  
Qy 2767 CAATCTGAAGCGGAGTATTTGTTTCATCGCATCAGCGGATTTTGGCCGCGCATCTTTACGC 2826  
Db 4021 CCATGTGAACAGTGGCTACTACTGCGTGACCTCCCGCATTTTCGCGCGGATCTCCGCA 4080  
Qy 2827 GGCGATGGTCAACCATCTGTGACAAACACAGCGGTACACAGGTAAACCCACGAGT 2886  
Db 4081 ACAGTTTGGACAAATTTATGGTGCATACCGCGTTTCAGCGCTTCATGCTGAATCGGT 4140  
Qy 2887 GACGTGGCGGATGGCGTGAACCTTGTGCGCAAGTGTGATTTGATGCTCGCGGCTGCA 2946  
Db 4141 CCAGTTAGCGATGGCGGATTTATTCAGCCAGTACAGTATGACGAGCGGTTTACAC 4200  
Qy 2947 GCGAGCGCAATCTGAGCTGGGTATTCAGGTGTTTCTTGGACAAAGATGTCAGCTGC 3006  
Db 4201 GCCTGATTTCTGCACTACGCGTAGGATTCAGGCAATTTATCGTCAAGAGTGGCAACTGAG 4260  
Qy 3007 GCAGCGCACGCGCTGACGAGCGCATCTGATGGATGCCACCGTGCATCAGCAAGCGG 3066  
Db 4261 GCGCGCATGGTTTATCGTCAACCGATTCATGGATGCGACGCTGCATCAGCAAAATGG 4320  
Qy 3067 TTATCGTTTGTCTACAGCTGCGCTCAGCGCGCATCGGCTATTTGATTTGAAGTACCCA 3126  
Db 4321 CTACGCTTTGTTTATACCTGCGCTTTTCGCAACCGCATCTGCTGATCGAGACACACA 4380  
Qy 3127 TTAAGTTAAACAGCGCGCTGGCGGAGAACACCGCTGTCAGCAACATGCGCGCATGTC 3186  
Db 4381 CTACATTTGACAAGGCTATCTTCAGGCGAAGCGCGCTCAGAACATTTTCGCGATATGC 4440  
Qy 3187 CAATCAGCAAGCTGAGCGCTGAGTACGCTGCTGCTGAAGACGCGCATATTTACCGAT 3246  
Db 4441 TGCGCGACAGGTTGGCGTTTACAGCGTTTCTGCGGGAAGAACAGGTTGATTTGCCAT 4500  
Qy 3247 TACCTGAGCGCGCAACATCGATTCGATCTGCAACAGCAGCGCGCGCAAGCTGACGCG 3306  
Db 4501 TACGTTAACGGCGATATCTGTCAGTTTGGCAACAGCAAC---CGCAAGCTGTAGCGG 4557  
Qy 3307 CCGCGCGCGGCTGTTTCATGCGAACACCGGTTTCTTTCGCGTCCGCGTGGCGCT 3366  
Db 4558 ATTACGCGCGGCTGTTTCATTCGCAACCGGCTACTCCCTACCGCTCGCGTGGCGCT 4617  
Qy 3367 AGCGGATTTGATGAGCGCTGTTGCCCAACCGATGCTCAGCTCAGCGCAACATATCGA 3426  
Db 4618 GGCGCATCTCAGCGCTGGATGTTTACCTTCTTCTGTTTACCAAGCATTTGC 4677  
Qy 3427 ACCTTTTCCCGCTCAGCAGTGGCGGAAACAGCGATTTTTCGCTGCTGCTAAACCGCATGCT 3486  
Db 4678 TCATTTTCCAGCAACGTTGGCAGCAACAGGTTTTCGCGATGCTGATCGATGTT 4737  
Qy 3487 GTTTTGGCGGTAAAGCGCAGCAGCGCTGCGGTGATGCAACGTTTTTACCGCTCGA 3546  
Db 4738 GTTTTGGCGGACCGCGCGAGTACGCTGGCGTGTGATGACGCTTCTATGGCTTACC 4797  
Qy 3547 TGCCGGTTTATAGCGCTTTTACCGCGGCACTGCGCTGCGCGATTAACCGGAT 3606  
Db 4798 CGAGATTTGATGCGCGCTTTTATGCGGGAATCTACCGTACCGATCGGCTACGAT 4857  
Qy 3607 TCTGCGCGCAAGCGCGCTGCCCATCGGTGAAGCGCTGCGCGCTGTTGAAATCTGCT 3666  
Db 4858 TCTGAGCGGCAAGCGCGCTTCCCGTTTCCGGCATTTGAGGCAATTTATGACGACTCA 4917



QY 3667 CGAACCGGAGAAATAATGAACGCACTTATGTGATTGGCGCAGGCTTTGGGGCGCTG 3726  
DB 4918 TCGTTGAAGAGCGACTACATGAACCAACTACGGTAATTGTGTGGGCTTTTGTGGGCTG 4977  
QY 3727 GCGCTGGCGAATTCGCTCTCAAGCGGCGGCATACCAACCACTTACTCGAGCAGCGGAC 3786  
DB 4978 GCACTGGCAATTCGTTTACAGCGCGCAGTATCTGTTTGTCTTGAGCAGCGGAC 5037  
QY 3787 AAACCGGCGGACGGGCGCTATGTGTGTGAGACAGTGGCTTTACCTTCGATGCGGACCC 3846  
DB 5038 AAGCGGCTGGCGGCGCTATGTGTATCAGGAGCAGGCGCTTTACTTTTGTATGCAAGGCGCT 5097  
QY 3847 ACGGTGATCACCGATCCCGAGCGCATCGAAGATTGTTCACTGGCAGGAAATCGCTC 3906  
DB 5098 ACCGTTATCACCGATCCCGAGCGGATGAGAACTGTTTGTCTTGGCGGTAAACGCTT 5157  
QY 3907 AGCGATTACGTAGCTGATCGCGGTAAAGCGCTTCTATGCGCTGTGTGGGAAGATGGC 3966  
DB 5158 AAGGATTAGCTGAGCTGTGCGGCTCACGCGTATTTATCGCTGTGTGGGAGTCCGGC 5217  
QY 3967 AAACAGCTTATTACGACAAATATCAGCGGCTGTGAGCAGCAGATCGCCACGTTCAAT 4026  
DB 5218 AAGGCTTCAATTAAGATAACACAGCGCCAGTTAGAAAGCGCAGATACAGCAGTTTAA 5277  
QY 4027 CCGCAAGATGTAGAAGGCTATCGTCAATTTCTTGCTTATTCACGTGAAGTATTTAGAG 4086  
DB 5278 CCGCGGATGTGCGGTTATCGAGGCTTCTTGACTATTCGCGTGGCGGTATCAATGAG 5337  
QY 4087 GGTATCTGAATCTCGGACGGTGGCGTTTCTGAGGTCGTGACATGCTGGCGCTCGC 4146  
DB 5338 GGTATCTGAATCTCGGACGGTGGCGTTTCTGAGTATTCGCGTGGCGGTATCAATGAG 5397  
QY 4147 CGCAGTTGGGAGCTGCAACATGCGGCGCTACAGCATGCTGGCGAAATTTAT 4206  
DB 5398 CCCAGTTGGCAAGCTGCGAGCATGCGCAGCGTTTACAGTAAGTTTCCGCGTACAT 5457  
QY 4207 CAGGAGCATCTGCGTCAGCGGTTTCTTCCATCTCATTTGCTCGTGGGCGTAACTCT 4266  
DB 5458 GAGGATGACATCTTCCGCGAGCGTTTCTTTTCACTCGCTCTTATGCGGGGGAATCG 5517  
QY 4267 TTTGCAAGCTCATGATCTATACCTTAACTATGCTGCGTGGAGCTGAATGGGCGTGG 4326  
DB 5518 TTTGCAAGCTCTGCTCAATTTATACGCTGATTCAAGGTTAGAACGGGAATGGGCGTGG 5577  
QY 4327 TTTCCGCGGGGACCGCGCTGCTGTCAGGCGATGCGCGACTGTTTCAGGACTTG 4386  
DB 5578 TTTCCAGCGGTGGAAACCGGTGGCTGGTCAATGSCATGATCAAGCTGTTTCAGGATCTG 5637  
QY 4387 GCGCGGAGCTGTTACTGAATGCGGAAGTGAGCAGCTGGAAACCGAGCGCAATCGCAT 4446  
DB 5638 GCGCGGGAAGTGTGCTTAAACGCGCGGTCACTCATATGAAACCGTTGGGACAAAGATT 5697  
QY 4447 AGCGGCTTCAGTTAGGCGGACACGCTTCGATGCGCGCTGTGGCTCCCAATGCC 4506  
DB 5698 CAGGCGGTGAGTGTGAAGACGCGACCGGTTTGAACCTGCGCGTGGCTGGAAGCT 5757  
QY 4507 GAGCTGGTGATACCTACGACAACTGCTTTCGCAACCATCGCTGGCAATGAAGTGG 4566  
DB 5758 GATGTTGTACATACCTATCGGATCTGCTGTCTCAGCATCCCGCAGCCGCTAAGCAGG 5817  
QY 4567 ACATCGTGAAGCGTGAAGCGCATGAGCACTCGCTGTTGTACTCTATTTTGGCTGTAAT 4626  
DB 5818 AAAAACTGCAATCCCAAGCGTATGAGTAATCACTCTGTTGTACTCTATTTTGTCTCAAC 5877  
QY 4627 CAGCGCATGAACAGCTCGCGACACACCGCTCTGTTTGGCGCGCTTATCGTGAATTG 4686  
DB 5878 CATCATCAGCATCACTCGCCCATATACCGTCTGTTTGGGCGCACGCTACCGTGAATG 5937  
QY 4687 ATCGATGAGATTTTCAACAGCAGCAGCTGCGCAGACGATTTTTCATTTACCTTGCACGG 4746  
DB 5938 ATTCAGGAATTTTAAACCATGATGCTGCTGAGGATTTTTCGCTTATTTTACACGCA 5997  
QY 4747 CCCTGACGACGATCCGTCTGCTGGCACCGCGGCTGGCGCAGCTTTTATGTGTAGCG 4806

DB 5998 CCTGTGTACGGATCCGTCACTGGCACCGGAAGGTGGCGAGCTATTATGTCTGGCG 6057  
QY 4807 CCGTGGCGCATCTCGGCACCGCTGACATCGATCGCAACAGGAAGAGCCGCGCTTGGC 4866  
DB 6058 CTTGTTCCACACTTAGCGACCGGCAACCTGATCGGCGGTAGAGAGCCCGACTGGC 6117  
QY 4867 GATCGAATTTTTCCTTATCTGAGCAGCACTACATCCCGGATTCAGTCAAGCAATAGTG 4926  
DB 6118 GATCGTATTTTTCGACTACCTTGAGCAACATTAATGCTGGCTTGGGAAGCCAGTTGGTG 6177  
QY 4927 ACACACAGATTTTACCGCGTTTGTATTTGCGCAACGCTGATGATGATGATGATGATG 4986  
DB 6178 ACGCACGATTTTACCGCGTTTGTATTTCCGCAAGAGCTCAATGCTGGCAAGGTTTCG 6237  
QY 4987 GCGTTTTGCTGAGCGGATTTTGAAGCAAGCGCTGTTCCGCGCATATAACCGGAT 5046  
DB 6238 GCTTCTCGGTTGAACCTATTTGACCCAGAGCGCTGTTTCCGACCAACATACCGGAT 6297  
QY 5047 GCGGATATCAGCAATCTTATCTGTTGGTGGTCCGTTACGATCCAGGCGGGGTTGCC 5106  
DB 6298 AAGCACATTTGATAATCTTATCTGTTGGCGCAGGCAACCATCTTGGCGGGCAATTTCC 6357  
QY 5107 GCGTGTATCGTTGCGCAAGCCACCGCAGGCTGATGCTGGAGATTCGCGCAATGA 5166  
DB 6358 GCGGTAATCGGCTCGGGAAGGCGACGCGAGGCTTAAATGCTGGAGGACCTGATTTGAGA 6417  
QY 5167 ATCGACAGCTTTTACTTTGAGCAAGTTAAACGCAACCATTGGCGGTGGCTCGAAGATTTCG 5226  
DB 6418 AT---AGCTATTACTGAATCATCGCTCGAACCATTGGCGTTGGCTCGAAGAGCTTTG 6474  
QY 5227 CCACCGCGCAAGCTGTTTGTATGACCGAGCGCGCCAGCAGCTGATGCTGTATGCT 5286  
DB 6475 CGACTGATCGACGCTTTTCGACCGCAAAACCGCTCGCAGCTGCTGATGCTTTTACGAT 6534  
QY 5287 GGTGCTCTCATGCGATGATGATTGATGATGGCAAAACGCTGGGCGAAGGCGGACGAGC 5346  
DB 6535 GGTGCGGCGCATGCGAGCGATCATTTGACGATCAAAACACTTGGGCTTTTCATGCCACGAG 6594  
QY 5347 ATGCGCTGGAAGACGCGCAGCAGCTATGACGATCTGCAAAATTTGAAACCGCGCGCT 5406  
DB 6595 CCTCTTCGAGATCCCTGAGCAGCGCTGACGAGCTTGAATGAAACCGCTCAGGCT 6654  
QY 5407 ACAGGCGCGCATGATGATGAAACCGCGGTTTGGGCGTTTACGAAAGTGGCAATCTC 5466  
DB 6655 ACGCGGTTTCGCAATGACAGAGCCGCTTTGCGCGTTTTCAGAGGTCGCGATGGCG 6714  
QY 5467 ACCAGCTGCGCAACAACTGGCGTTTGTATCATCTGGAAGGCTTCGCTATGATGACGCA 5526  
DB 6715 ATGATATCGCTCCCGCTTACGCGTTTCGACCATCTGGAAGGTTTTCGCAATGTCGCG 6774  
QY 5527 ACGAATATTACGAGCTTCGATGACAGCTGCGTTTACTGCTATCAGTTCGCGGCGTGG 5586  
DB 6775 AAAACGCTACTGACACTGACGATAGCTGCGTTTATGCTATCAGCTCGCGGCTGTTG 6834  
QY 5587 TCGGTTTTCATGATGCGCGCTTAATGGGCTGCGGACGAGCGGTCTCGATCACGCT 5646  
DB 6835 TGGGCTGATGATGCGCAATTTATGGGCTTTCGCAATGCGGCAACCGCACGCTCGATCGGCT 6894  
QY 5647 GCGATTTAGGACTGGCGTTTCCAGCTCACTAACTTGGCGCGCAATTTGTAGAGATGCG 5706  
DB 6895 GCGATCTCGGCTGGCTTTTCCAGTTTGAACCAATGCGCGTATTTTTCGACGATGCTC 6954  
QY 5707 AAAATGCTGCTGCTATCTGCGCAATCTGCTGATCAGCGGATTTACGCGCGGATA 5766  
DB 6955 AGGTGGGCGCTGTTTATCTGCTGAAAGCTGGCTGGAAGAGAGGACTGACGAAGCGA 7014  
QY 5767 CGCTGACTGCAACGCAACATGTCGAGCGCTCGCTCACTGGCAGCGGCTTTTGTGGCG 5826  
DB 7015 ATTATGCTGCGCAGAAACCGGAGGCTTTAAGCGGCTTTCGCGGCGACTGTTACGG 7074  
QY 5827 AGCGGAAACCTTATCATCTGCGGCGGATCCGTTTACCGGTTTACCGCTGCGCTCGG 5886



Db 7075 AAGCGGAACCTATTACGTATCATCAATGCCGGTCTGSCACAAATTACCTTTACGCTCGG 7134  
Qy 5887 CGTGGCCATCGTACGGCTCGCGGCTTTATCGCAAAATTTGGCGTCAAAAGTTTCAGCAG 5946  
Db 7135 CTTGGGCCATCGCGACAGCGAAGACAGTACCGTAAATTTGGCGTGAAGTTGAACAGG 7194  
Qy 5947 CGGCTGTGACGCGCTGGGATTCACGCGACGACGACCAAGTAAAGTGAAGAACTGGCGCTGC 6006  
Db 7195 CCGTAAAGCGGCTCGGATCATCGCCAGTCAACGCTCCACCGCGGAAATTAAGCGCTTT 7254  
Qy 6007 TGGTAAAGGGGAGGTTTGGCGATCACTTCGCGTGTCTCGTCTCGAACCAGCGCTCCGG 6066  
Db 7255 TGTGACGCGCATCGGTCAGGAGTTTACTTCCCGGATGAAGCATATCCACCCGCTCTG 7314  
Qy 6067 CTGCTGTGGAGCGTCTCGTTGATTTTATGTCGTCGACGCTGGCGGAGCGTGGCTTG 6126  
Db 7315 CTCATCTCTGGAGCGCC- - - - -GATCTAGCGCATGCCCTTCTCTCAGCGTGGCCTG 7368  
Qy 6127 CAGCTTATTTCAGCGTGGCGCTAGAGGAACCAACAGCACGCGCTTCACGCCCGG 6186  
Db 7369 AAGTTTAGATAACGGTGGCGGTACAGAAACCAAGGACACGCGCCCTCTTTCCCT 7428  
Qy 6187 CACCGCATGATGCGGTGGCCCATGTATAGCGCTTTAAGATAGCCTTTGCGGGGAT 6246  
Db 7429 TACAGCATGATGATACGCTGGGCCATGTATAACCGTTTCAGGTAGCCTTTGGCGGTAT 7488  
Qy 6247 ATAGCGGAACGCGAGCTGTATGACACAGGCCATCTGTGCAACATGAAGTAGAGCGCGC 6306  
Db 7489 GTAGCGGAACGCGAGCGCTGGTGTACCACTCGTGGGACCAATAAATACAGTAAACC 7548  
Qy 6307 GTACGTGCTATTCGCGACCAATCACTGACGCGCCACATGCTTGCACACGACATA 6366  
Db 7549 ATAAGCGTTCATCGCTGCAACCAATCCACTGGAGCGGCCAGATTCCTGTACTCCGAAGTA 7608  
Qy 6367 AATCAGCACATCGCCAGTACCGCAAAACCAACCGCATAAAGATCGTTGAGCTCAAACTT 6426  
Db 7609 ATCAGGGCAATCGACACATGGCGNATACCGCATAGAGATCGTTAACTTCAATGC 7668  
Qy 6427 ACCGTGTGCGTTTCAATGTGCGACAGATGCCAGCGCCATCCCCAACCGTGCATGATGA 6486  
Db 7669 GCCTTTACCGGTTTCATGATGTGAAGATGCCAGCGCCCAACCCAGCGTGCATGATGA 7728  
Qy 6487 TTTATCGACAGCGCGCTACGATTCATTCATCCATCCACCGCTTGCACAAAGATAGCAC 6546  
Db 7729 TTTATGTGCGAGTGAACCACTTCCATGCGGACCGCTGACCAACACCATCAGGCG 7788  
Qy 6547 GTTCCATACACAGAGCAATTTGCTGCCA 6574  
Db 7789 ATTCCAAATCCACACATATTTCTCAA 7816

RESULT 4

US-10-734-778-40  
; Sequence 40, Application US/10734778  
; Publication No. US20040209365A1  
; GENERAL INFORMATION:  
; APPLICANT: E.I. duPont de Nemours and Company, Inc.  
; APPLICANT: Suh, Wonchul  
; APPLICANT: Rouviere, Pierre  
; TITLE OF INVENTION: PARALLEL CHROMOSOMAL STACKING OF TRAITS IN BACTERIA  
; FILE REFERENCE: CL2026 US NA  
; CURRENT APPLICATION NUMBER: US/10/734,778  
; PRIORITY FILING DATE: 2003-12-12  
; PRIOR FILING DATE: 2003-12-12  
; PRIOR FILING DATE: 2002-12-19  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 40  
; TYPE: DNA  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Plasmid pPCB15

US-10-734-778-40  
Query Match 37.4%; Score 2619.6; DB 18; Length 8609;  
Best Local Similarity 64.7%; Pred. No. 0;  
Matches 4015; Conservative 0; Mismatches 2169; Indels 24; Gaps 7;  
Qy 373 CAGCTGCGACCAATTTTACAGCGCATCTTGAACATTTACTGCTGCCGACGACGAAAGC 432  
Db 1627 CAGTTGCTGGCTGATATCGATAGCCGCTTGATCATAGTTACTGCCGTTTCAAGGTGAGCGG 1686  
Qy 433 GATCCGCTGCGTCCGCGATCGTCCGGAACGCTGGCGCAGGCAAAAGTATTTGCTCTC 492  
Db 1687 GATTTGTGGTGGCGCGATCGTGAAGCAGCTGSCACCGGCAAAAGTATTTGCTCG 1746  
Qy 493 TTATTTACTGCTGCTGGCAGCGCGCATATGGGTGGAGCTGACGCAAAATGGCGTTCTC 552  
Db 1747 ATGCTGCTGTTATTAAACAGCGCGCATCTTGGCTGTGCATCAGTCAACGCGGATTTACTG 1806  
Qy 553 GATCTCGCTGTCAGTGGAAATGTCGACCGGCACTGCTGATTTCTGATGACATTTCC 612  
Db 1807 GATTTAGCTGCGCGTTGAAATGGTGCATGCTGCTGCTGATTTCTGATGATATGCC 1866  
Qy 613 TCGATGGATTAACCGCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 672  
Db 1867 TCGATGGAGATGCGCGATGCTGCGGGCGTCCACCATTTCAACGAGTACGCTGAA 1926  
Qy 673 AACGTGGCGATTTCTCGCGCCATCGCTGCTTACGCGCGCATTTTGAAGTATGTCATT 732  
Db 1927 CATGTGGCGATTTCTGGCGCGTCTTACTCAGCAAAAGCTTTGGGGTGTATGGCGAG 1986  
Qy 733 GCACCGGTTTGGCTGCGCATATCAATAAATCTGAAGCGATTTGCTGAACTCTCCGCTGCCGTC 792  
Db 1987 GCTGAAGGTCTGACCGCATAGCCAAACTTCGCGCGTGTGCGAGCTGTCTCACTGCGATT 2046  
Qy 793 GCGCTGCGAGGCTTAGTCAAGGGCAATTTCCAGGATCTGACGACGCGACGACGAGCCCGC 852  
Db 2047 GCGATGCGAGGCTGCTGTTGAGGCCAGTTTAAAGACCTCTCGAAGCGCATTAACCCCGC 2106  
Qy 853 AGCCCGAAGCGATGCCATGACCAAGAACTGAAACACAGCGTCTGTTTCCGCGCACG 912  
Db 2107 AGCGCGGATGCCATCTGCTAACCAATCAGTTTAAACCCAGCACGCTGTTTGGCGGTCA 2166  
Qy 913 CTGCAATGGCGCATTTGCCGCTGACGCTTACCGAGGTCGCGGCAAGACTTAGCTTC 972  
Db 2167 AGCAAAATGGCGTCTCATTTGCGGCCAACCGCTCTCGAAGCGCTGAGAACCTGCACTCGT 2226  
Qy 973 TTGCGCCAGGATTTGGCGCAGCGCTTTCAACTGCTCGACGCTCGCGACGCTTGCATA 1032  
Db 2227 TTCTGCTGATCTCGCGCAGCGCTTTCACTGCTGACGATCTTACCGATGCGATGACC 2286  
Qy 1033 CACACCGTAAAGATGTGCAACAGATCAGGGCAATCAACGCTGGTACAGATGCTCGGT 1092  
Db 2287 GATACCGGCAAGACATCAATCAGGATGAGGTAAATCAACGCTGCTCAATTTATTTAGGC 2346  
Qy 1093 GCTGACGGCGGAAAGCTGCGCTGCGCATCACTGCGCAGCGCAGATGACACCTTGC 1152  
Db 2347 TCAGCGCGGTCGAAGAAACGCTGCGCAGCATTTTGGCGCTTGGCCAGTGAACACCTTTCC 2406  
Qy 1153 TCGCGCTGCTCATCGCGCATCGCCACTCGCCAA- - -TATATGACGCGCTGTTTAAATCAA 1209  
Db 2407 GCGGATGCAAAACGCGCATTTCCACCACTTTTATTTTATTTTATTTTATTTTATTTTATTT 2466  
Qy 1210 CAGCTAGCGATATTTCAACTGAGCGCGGCTCAGCGCGTGGGCACTTTTGGCGTATGCGC 1269  
Db 2467 AAACCTGCTGCGCTCAGTTAAGATGCTGCA- - -TGAGCCATTTTGGCGTATGCGAC 2521  
Qy 1270 CGCGCTCTACAGCCACTTTACGCGTTGCGAGCGTTAGCAAAACGCTGCTGGCGCG 1329  
Db 2522 CGCCCTTTTTCAGCCATGTTGCGCTCTGCAAAACCTTGTCTCAGGAATTTAGTGGCGCGG 2581  
Qy 1330 GCGATCGCATCATTTTCATCCAGCAAGCGGATGCGCGCACTTTTCTTACGACGACGCA 1389  
Db 2582 GTATCGGTACGTTTTTTTACGCAACATGACTGCAAAAGCGCTGTTACCGGACGCGATA 2641

QY 1390 TCGATTTTGTGCGCTCGGCGCAACAGACGATTCCTGCGGTTCCCTGCGCGCCCGCTGTC 1449  
 Db 2642 TCGGATTCAGACCGTCGCACTGCAACGCACTCTCCGCGTTCTTATCGCACCTGCTGC 2701  
 QY 1450 ATCGGCTGGCTCGCGCGCGCGCTGTCGCTGTTTTCGGTGATCGACGATCTCGCGTCT 1509  
 Db 2702 ACCCTGGCGCGCACCACTCGGACCTCGATGTTTACGACTGATCAATGAATGSCAGTA 2761  
 QY 1510 GCACCGATATGCTGTCGCGCAACTGCTCGGTA CTGAAAGCAATGAAATCGATGGCG 1569  
 Db 2762 CCAGCGATATGCTTTCGCGGAACCTGCGCGCTTTTCATGCGTTGAGATAGAGGCG 2821  
 QY 1570 TGATCGCGACGAATGGAAGCGCGCGGAGTGGTCTGCTGAGAGCGCTGCATCTCGCT 1629  
 Db 2822 TGATCGTTGATCAATGGAAGCGCGGAGTGCAGTAGTCGCAAGAGCTGACGCTCGCGCT 2881  
 QY 1630 TTGTTTTCGCTGCGCTTTCGCGCTCAATTCGTGAAGCGCGGATTCGCTTTCGCTGA 1689  
 Db 2882 TTGTTTTCGCTGCGCTTTCGCGCTCAACCGCGGACCGGTTTTCCTCTGCGCGTGA 2941  
 QY 1690 TGCCCTTCGTTTTCGACAGGATGA CAAGCGCTGAAAGCTTTTCAGGCCAGCAGGATA 1749  
 Db 2942 TGCCCTTCGATACGCGACCGGATGCGGCTCGGGAAGCTATACCAACGAGCAAAAA 3001  
 QY 1750 TCTATGATCGCATCATGCTGCTGCGGCGACGTGATCTCAACACAGCGCGGGGTTTA 1809  
 Db 3002 TTTATGACTGGCTGATGCGACGTGACGATCGTGTGATCGCGATCATGATGCGAATGG 3061  
 QY 1810 ATTGACGAGGCGCGCGGATACATCAGTGTCTGCGCGCTGCGCA CAATACGCGAGA 1869  
 Db 3062 GTTTAGCCCGGTGAAABACTGCATCATTTGTTTCTCCACTGGCA CAATACGCGAGT 3121  
 QY 1870 TGCTGCGGCTTTGATTTTTCAGTGCAGCACTGCGCGCTGCTATCAGCGCGTGGGCG 1929  
 Db 3122 TGA TCCCGCAACTGAGTATTTCCCGCAAGCGCTGCCAGACTGCTTTTCATGCGGTGAC 3181  
 QY 1930 CACTTCGCG - - - CCGCGTTTCTCTGCGCGCTCCATGCGCGCTGCGCGAGCGTGCCTC 1986  
 Db 3182 CGTTACGCGAACCCGAGGAGCGCGCGGTTCATCAACTTCTTATTTTCGTCGCCGACA 3241  
 QY 1987 AGCGGTGTTTATGCTGCTGGTACGTTGAGGCGATCGCTTCGCGCTGCTGCTGCTGCT 2046  
 Db 3242 AACCCCGTATTTTGTGCTGCTGCGGCACTGCGAGGACATCGTTATGCGCTGTTTCAGGA 3301  
 QY 2047 ATCTGCGCAGCGCTGCGCGCTGCGGCTATCGCTGCTGATGCGGCTGCGCGAGCGTGC 2106  
 Db 3302 CCATGCGCAAGCGCTGGAAGAGTGA TCGCGAGTTACTGTTGGCACACTGTTGGCGGC 3361  
 QY 2107 TAAACGCGCAACAGACGATCAGCTGAGCTCGCTGCGCGCGCTGCGGTGACGGAATTCG 2166  
 Db 3362 TCTCAGCCACGACGAGGAGTGAATGSCCGCGCGGAGACATTCAGGTTGTGGAATTTG 3421  
 QY 2167 TCGATCAGCGCGAGCGCTCAGACGCGGAGCTGTTTATCATCTCATGCGGTTAAACA 2226  
 Db 3422 CCGATCAATCCGAGCACTTTTCAGGCAAGTTGA CAATCAACATGTTGGATGAATA 3481  
 QY 2227 GCGCGTGAAGCACTGGAATGCGTACGCGGATGCTGCGCTGCGGCTGCGGATTTGATC 2286  
 Db 3482 CGGTACTGAGCGTATTTGCTTCCGCAACGCTACTGCGCTGCGGCTGCGGCTGCGGATTTGATC 3541  
 QY 2287 AGCCCGCGTGGCGCGCGATGAGTGGGATGAGCGTTGCTGCGCGCGCATCAGCTTTA 2346  
 Db 3542 AACCTGGCGTGGCATCAAGAAATTTTATCATGCGATCGGCAAGCGTGGCTCTCGGTTA 3601  
 QY 2347 GCGGTGTTTCATCACTGGAGCAGCATCTGCAACAGCTGCTGACGAGCATCGTTACGCGC 2406  
 Db 3602 CTACAGCAGCTGCGCTGCGCGGAGATTCGATCGCTGCTGATCAACCGATTAACCGC 3661  
 QY 2407 TACGGAATGTCAGGATTCAGGCGAGCTGACAGCGCGAGCGGTTGCGAGGCTGCGCGC 2466  
 Db 3662 AGCGTATGACAAAAATTCAGGCGCGATTTGCGTCTGCGAGGCGGACACCGAGCGCGCGC 3721

QY 2467 ACATCGTCGAGCAGCGCTGTGCCAGCAAGTGTGCTGCGGAGGCGGACCTGATGCG 2526  
 Db 3722 ATATTTCTGAACAGGCGATGCG - GACCTGTACGCACTGACTCAGTGGGAGGATTTATGCA 3780  
 QY 2527 CAGCGAATACGATGATTTTGGTGGTCTGCACTGGGCAATGGCTTCAATGCGCTGCG 2586  
 Db 3781 ACCGCACTATGATCTCATTTCTGGTGGTCCCGCTTGGCTTAATGGCTTTATCGCGCTCG 3840  
 QY 2587 TCTGCGCTCAATTCAGCCCAACTGAAATGCTCTGTTGCTGAGAGCGATGCGCATCCGCG 2646  
 Db 3841 GCTTCAGCAACAGCATCCGATATCGGATCTTGCTTATTGAGCGGGTCTCTAGGCGCG 3900  
 QY 2647 AGCAATCATPACTGCTGCTTCATCAGAGCGATCTCAGCGCGGCAACAATTCGCTGGCT 2706  
 Db 3901 AGGAACCATACCTGCTCTCTTTCAGAAAGAGATTTAACTGATCAAGCATCGCTGGAT 3960  
 QY 2707 GCAACCGCTGATTAACGCTGTTGGTGTATCAGGTGCGTTCCTGCTGCGCTGCGCG 2766  
 Db 3961 AGCGCGCTTGTGGTTCATCTGCGCGGCTACTGCGTACCTCCGCGCATTTCCGCGGAT 4020  
 QY 2767 CAATCTGAGCGGGATTTATTTCCATCGCATCAGGCGATTTTGGCGCGCATTTTACGC 2826  
 Db 4021 CCATGTGAACAGTGGCTTACTGCTGACCTCCGCGCATTTCCGCGGATTAATCCGCGCA 4080  
 QY 2827 GCGATGGGTGACGATCTGTGGACAACACAGCGGTACAACAGSTAAACCCACGCGAGT 2886  
 Db 4081 ACAATTTGGAACAATTTATGGTTCATACCGCGTTCAGCCGTTTTCAGCCGTTCTGCTGAATCGGT 4140  
 QY 2887 GAGCTCGCGGATGGCGGTGAACCTTGTGCGCAAGTGTGATTTGATGCTGCGCGCTGCA 2946  
 Db 4141 CCAGTTAGCGGATGGCGGATTTATTCATGCGAGTACAGTGCATCGACGAGCGGTTACAC 4200  
 QY 2947 GCGACGCCACATCTGAGCTGGGTATCAGGTGTTTCTTGGACAGAGTGGCAGCTGGC 3006  
 Db 4201 GCCTGATCTGCACTAGCGGTAGGATTTCCAGGCAATTTATCGGTGAGGTTGGCAACTGAG 4260  
 QY 3007 GCAGCGCAGCGCTGCGAGCGGATCTGATGATGCGCACCGTCCGATCAGCAAGCGG 3066  
 Db 4261 CGCGCGGATGGTTTATCTGTCACGATTAATGATGATGCGCGCTGATCAGCAAAATGG 4320  
 QY 3067 TTATCGTTTGTCTACAGCTGCGCTGAGCGCGATGCGGTATTTGATGAGATACCA 3126  
 Db 4321 CTACCGCTTTGTTTATACCTGCGCTTTCGCAACCGCACTGCTGATCGAAGACACACA 4380  
 QY 3127 TTACGTTAACAGCGCGCTGCGGAGACACCGCTGCTCAGCACATCGCCGCTATGCG 3186  
 Db 4381 CTACATTCAGCAAGGCTAATCTTCAGGCGGAAGCGCGGTGAGAACATTCGCGATTATGC 4440  
 QY 3187 CAATCAGCAAGGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3246  
 Db 4441 TCGCGGACAGGTTGGCGTTACAGAGCTTCTGCGGGAAGAACAGGCTGCTGCTGCTGCT 4500  
 QY 3247 TACCTGAGCGGCAACATCGATTCGATTCGCAACAGCAGCGCGGCGGAGGCTGCGCG 3306  
 Db 4501 TACGTTAACGCGGATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4557  
 QY 3307 CTTGCGCGCGGCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3366  
 Db 4558 ATTAAGCGCGCGGCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4617  
 QY 3367 AGCGGATTTGAGCAGCGCTGTTGCCCAACGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 3426  
 Db 4618 GCGCGATCGTCTCAGCGCGCTGATGTTTACCTTCTTCTGCTGCTGCTGCTGCTGCTGCT 4677  
 QY 3427 AGCTTTGCGCTGAGAGTGGCGGACAGGATTTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCT 3486  
 Db 4678 TCACCTTTGCGGAGCAAGTGGCAGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4737  
 QY 3487 GTTTTGGCGGTAAGCGCGCAGCAGCGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3546  
 Db 4738 GTTTTGGCGGACCGCGGAGTACGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4797  
 QY 3547 TGCGCGGTTAATAGCGCGCTTTTTCAGCGGCGCACTGCGCGCTGCGGATTAACCGCGAT 3606

Db 4798 CGAGATTTGATTCGCGCTTTTATGCGGAAACTCAACCGTACCGGATCGGCTAGCAT 4857  
Qy TCTGTGCGGCAAGCCGCGGTGCCATCGGTGAAGCGCTGCGCGGCTGTGGAATCTGT 3666  
Db TCTGAGCGGCAAGCCGCGGTTCCTGTTTCGCGCATTCGAGGCAATATGACGACTCA 4917  
Qy CGAACGAGGGAAGAAATGAACGCACTTATGTGATTTGGGCGAGGCTTTGGCGGCTG 3726  
Db TCGTTGAAGAGGACTACATGAACCAACTACGGTAATTTGGTGGCGGCTTTGGTGGCTG 4977  
Qy GGGCTGGGCAATTCGCTGCAAGCGGCGGCATACCAACACCTTACTCGAGCAGCGGAC 3786  
Db GCACTGGCAATTCGTTACAGCGCGCAGGTATTCCTGTTTCTGCTGCTGAGCAGCGGAC 5037  
Qy AAACCGGGCGGACGCGCTTATGTGTTGAGGACAGTGGCTTTACCTTCGATCCGAGCCC 3846  
Db AAGCGGGTGGCGGCTTATGTATCAGGACGAGGCTTTACTTTTGTGATCGAGCCCT 5097  
Qy ACGGTGATCACCGATCCGAGCGGCATCGAAGAGTGTTCACGCTGGCAGGAAATCGCTC 3906  
Db ACCGTTATCACCGATCCGAGCGGATTTGAAGAACTGTTTGTCTGCGCGGTAAACAGCTT 5157  
Qy AGCGATTACGTCGAGCTGATCGGTAACGCCCTTCTATCGCTGTGCTGGGAAGATGSC 3966  
Db AAGGATCTGAGCTGTGCGGTGCGGTACCGCTTTTATCGCGTGTGCTGGAGTCCGCG 5217  
Qy AAACAGCTTTGATTACGCAATTAATCAGCGCGTGTGGAGCAGCATCGCCAGTTCAAT 4026  
Db AAGGCTCTCAATACGATAACACGAGCGCCAGTGTAGAGCGCAGATACAGCAGTTAAT 5277  
Qy CCGCAAGATGTAGAGGCTATCGTCAATTTCTTGTGCTATTCAGTGAAGTATTTAGAG 4086  
Db CCGCGCATGTTGCGGGTTATCGAGCGTTTCTTGACTATTCGCTGCGGTATTTCAATGAG 5337  
Qy GTTATCTGAACCTGGCAGCGTGCCTTTCTGCAAGTGGTGACATGCTGCGCTCGCG 4146  
Db GCTATCTGAAGCTGGCACTGTGCTTTTATGCTTCAAGACATGCTTCGGCGCGG 5397  
Qy CCGCAGTTGGGACGCTGCAAGCATGGCGACGCTTACAGCATGGTGGCGAAATTTATT 4206  
Db CCCAGTTGGCAAGCTGCAGCATGGCGACGGTTTACAGTAAAGTTGCGGCTACATT 5457  
Qy CAGGAGATCATCTGGTCAGCGGTTTCTCTTCCACTCATGCTGGTGGCGGTAACTCT 4266  
Db GAGGATGAGCATTTGCGGAGCGTTTCTCTTCTCACTGCTCTTATGTTGGGGGGAATCG 5517  
Qy TTTGCAAGTCTCATCTATACCTTAACTCATGCTGGAGCGTGAATGGGCGGTGG 4326  
Db TTTGCAACCTCTCATTTATACGCTGATTCAGCGTTAGAACGGGAATGGGCGGTCTGG 5577  
Qy TTTCCGCGCGGCAACCGCGCTGTGTGCGAGGCAATGGCGCACTGTTTCGAGGACTTG 4386  
Db TTTCCAGCGGGTGAACCGCTGGCTGGTCAATGCGCATGATCAAGCTGTTTCAGGATCTG 5637  
Qy GCGCGGAGCTGTTACTGAATCGGAAAGTGAAGCAGCTGGAAACCGAGCGCAATCGCAT 4446  
Db GCGCGGCAAGTGTGCTTAACGCCCGGGTCACTCATATGAAACCGGTTGGGGACAAGATT 5697  
Qy ACGCGGCTTCAGTTAGAGGGCGGACGACGCTTCGATGCGCGCTGTGSCCTCCAATGCC 4506  
Db CAGCGCGTTCAGTTGAAGCGCGACCGGTTTGAACCTGCGGCTGGCGTGGCGTCAACGCT 5757  
Qy GACGTGGTGCATACCTACGACAACTGCTTCGCCCACTCCGCTGGCAATGAAGTGGG 4566  
Db GATGTTGTACATACCTATCGCGATCTGCTGTCTCAGCATCCCGCAGCGCTAAGCAGGG 5817  
Qy ACATCGCTGAAGGTTAAGCGCATAGCAACTCGCTGTTGTACTCTATTTTGGCTGAT 4626  
Db AAAAATCTCAATCCAAAGCGTATGAGTAACTCACTGTTGTACTCTATTTTGTCTCAAC 5877  
Qy CAGCGCATGAACGCTCGCGCACACACGCTCTGTTTGGCGGCTTATCGTAGTTG 4686

Db 5878 CATCATCAGGATCAACTCGCCCATCATACCGTCTGTTTTGGGCCACGCTACCGTGAAC 5937  
Qy ATCGATGAGATTTTCAACAGCAGCAGCGCTGGCAGACGATTTTTTCACTTTACCTGACGCG 4746  
Db ATTCAGGAAATTTTAACTGATGCTGCTGAGGATTTTTTCGTTTTATTTACACGCA 5997  
Qy CCCTGACAGGAGTCCGTCGCTGCGACCGCCCGCTGGGCGAGCTTTATGTTGTAGCG 4806  
Db CTTGTGTCACGAGTCCGTCATCGCACCGGAAGGTTGGGAGCTTATTTATGTTGCTGGCG 6057  
Qy CCGGTGCGCATCTCGGCAACCGCTGACATCGACTGCAACAGAAAGGACCGGCTTGGCG 4866  
Db CTTGTTCCACACTTAAAGCAGCGGCACTCGACTGGGCGGTAGAAGGACCCGACTGCGC 6117  
Qy GATCGAATTTTGTCTTATCTGAGCAGCATCATGATGCGGGGATTCGTGAGCAATTTAGTG 4926  
Db GATCGATTTTGTGACTACTGTTGAGCAACATTTACATGCTGCTGGCTTGCAGAACCAAGTTG 6177  
Qy ACACACAGATGTTTACCGCGTTTGTATTTTGGGACACGCTGATGCCATCACGCTCG 4986  
Db ACACACCGTATGTTTACCGCGTTTGTATTTTGGGACACGCTGATGCCATCACGCTCG 6237  
Qy GCGTTTTGCTGGAGCGGATTTTGAACGAAAGCGCTGTTCCGCCCGCATAAACCGCGAT 5046  
Db GCTTCTCGTTGAACCTTATCTGACCGAGCGCTGTTCCGACCATTAACCGCGAT 5297  
Qy GCGGATACGCAATCTCTATCTGTTGGTGGCGGTAGCATCTCAGGCGCGGCGTCC 5106  
Db AAGCACAATGATTAATCTTTATCTGTTGGCGCAGCACCCATCTCTGGCGGCAATCC 6357  
Qy GCGGATGATCGGTTGGGCAACCGCACCGCTGATGCTGGAGGATCGCGCGCAATGA 5166  
Db GCGGTAATCGGCTCGGCGAAGCGCAGCGCATTAATGCTGGAGGACCTGATTTGACGA 6417  
Qy ATCGACAGCTTTTACTTGAAGTAACGCAAAACCATGCGCGTGGCTGCAAGAGTTTCG 5226  
Db AT---ACGTCATTGATTAATCATGCCGTGAAACCATGCGGTTGGCTCGAAAGCTTTG 6474  
Qy CACCGCGCGCAAGCTGTTGATGACCGCGCGCGCATGCTGATGCTGATGCT 5286  
Db CGATGCTACGAGCTTTTCGACGCCAAAACCGGTGCGAGCGTGTGATGCTTTACGAT 6534  
Qy GGTGCTGCTACGCGATGATGATGGCAACCGCTGGCGAAGCGGCGACGCGCAGC 5346  
Db GGTGCGGCACTGCGACAGCTCATTTGATGATGATGATGATGATGATGATGATGATGAT 6594  
Qy ATGCGTGAAGACGCGCAGCGCATGATGACGATCTGCAATTCGAAATTCGAACTGAGGCT 5406  
Db CCTCTTCCAGATGCTGAGCAGCGCTGACGAGCTTGAATGAAATGAAATGAAATGAAAT 6554  
Qy ACAGCGCGCGCATGATGAACCGGCGTTTGGGCGTTTACGAGGTTTGGGCGATCAATTC 5466  
Db ACGCGGTTGCAAAATGACAGCGCGCTTTTGGCGGTTTTCAGGAGTTCGCGATGCGCG 6714  
Qy ACCAGTCCCGCAACACTGGGCTTTGATCATCTGGAAGGCTTCGCTATGATGACGCA 5526  
Db ATGATATCGCTCCCGCTACGCGTTTCGACCATCTGGAAGGTTTGGCAATGATGATGATG 6774  
Qy ACGAATATTACGAGCTTCGATGACGCTGCTGCTTACTGCTTATCACTGCGCGGCGTGG 5586  
Db AAACCGGCTACCTGACATGAGACGATACGCTGCGTTATGCTTATCACTGCGCGGTTG 6834  
Qy TCGGTTTGAATGAGCGCGCTTAATGGGCTGCGCGACGAAAGCGGTGCTGATCAAGCTC 5646  
Db TGGGCTGATGATGCGCAAAATATGCGGCTTCGCGAATAACCCACGCTCGATCGCGCT 6894  
Qy GCGATTTAGGACTGCGGTTCCAGCTCACTAATTCGCGCGACATTTGTAAGATGCGG 5706  
Db GCGATCTCGGCTGCGCTTTTCCAGTTGACCAATTCGCGGCTGATTTGTCGACGATGCTC 6954  
Qy AAAATGTTGCTGCTTATCTGCGCAATCTGCTGCTGATCAGCGGATTTACGCGCGGATA 5766  
Db AGTGGCGCGCTGTTTATCTGCTGAAAGCTGCTGAAAGAGGAAAGGACTGACGAAAGCGA 7014



[illegible]

Db	3602	CTACCAAGCCATGCGCTGGCGCGGAGATTGATCGTGTCTGATCTAAACACGATTA	CCCCG	3666
Qy	2407	TACCGEATGTCAGCGATTTCAGCGCGCAGCTGACGCGCGCAGCGGTTCGCCACGCTG	CGCGCG	2466
Db	3662	AGCGTATGACAAAAATTCAGGCGCATTTGGTCTGGCAGGCGGCACACAGCGCGCGCG	3721	
Qy	2467	ACATCGTCGAGAGGCGCTGTGCCAGAGCAAGTGTGTGGCGGAGAGCGACTGATGCG	2526	
Db	3722	ATAATGTTGTAACAGGCGATCGC-GACCTGTTCAGCAGTACTCAAGTGGCAGGATTA	TCGA	3780
Qy	2527	CACGCAATACGATCTGATTTTGGTTCGGTCTGCACATGCGGAATGGCTTGAATTCG	CTGG	2586
Db	3781	ACCGCACTATGATCTCAATCTGGTTCGGTCCGGTCTGGCTAAATGGGCTTATCG	CTCG	3840
Qy	2587	TCTCGGTCAATTCGAGCCACAATGAAATGCCCTGTGCTGGAGAGGCGATCGCAT	CCGGC	2646
Db	3841	GCATTACGACATCCGATATCGGATCTTTGCTTATTTAGCGGGTCTCTAGCGGG	3900	
Qy	2647	AGGCAATCATCCTGTGCTTTCAATCAGCGATCTCAGCGCCGAAACAATTCCTGTG	CGCT	2706
Db	3901	AGGNAACATACCTGGTCCCTTTTCAGAAAGAGATTTAAACGCTGAATCAGCAT	CGCTGGAT	3960
Qy	2707	GCAACCGCTGATTACCGTGGCTTGGTCAGGTTATCAGTTCGTTTCTCGCTGCGCG	2766	
Db	3961	AGCGCCGCTTGTGGTCCATCACTGGCCGACTACCAAGTTTCGTTTCCCCAAC	CGCGTCG	4020
Qy	2767	CAATCTGGAAGGGGATTAATGTTTCCATCGCATCAGCGGATTTTCCCGCCCAT	CTTTACGC	2826
Db	4021	CCATGTGAACAGTGGCTACTACTCGGTGACTCCCGGCATTTCCCGGGATACT	CCGGCA	4080
Qy	2827	GGCGATGGGTGACGATCTGTGGACAAACACAGCCGTACAACAGGTAAAAACCA	CGAGGT	2886
Db	4081	ACAGTTTGGACAAACATTTATGGCTGCATACCGGGTTTCAGCCGTTTCATGCT	GAATCGGT	4140
Qy	2887	GACCTGGCGGATGGCGGTGAACTTGCTGGCAAGTGGTGAATGTGTCGGCGCTG	CA	2946
Db	4141	CCAGTTAGCGGATGGCGGATTAATCATGCCAGTACAGTGATCGACGAGCGGGT	TACAC	4200
Qy	2947	GCCGACGCCACATCTGCAGCTGGGTTATCAGGTGTTTCTTGGAACAAGATGG	CAGTGGC	3006
Db	4201	GCTTGATTCGCACTACGCGTAGATTCAGGCAATTTATCGGTACAGAGTGGCA	CTGAG	4260
Qy	3007	GCAGCGCACCGGCTGCAGACGCGGATCTTGATGGATGCCCGTGCATCAGCA	AGCGGG	3066
Db	4261	CGCGCGCATGGTTTATCGTCACCGATATCATGTGATCGACGTCGATCAGCA	AAATGG	4320
Qy	3067	TTATCGTTTGTCTACAGCTGCGCTCAGCGCGGATCGGCTATTGATTTGAAGAT	ACCCA	3126
Db	4321	CTACCGGTTTGTTTATACCTGCGGCTTTCGGCAACCGCATCTGCTGATCGA	AGACACA	4380
Qy	3127	TTACGTTAAACAGCCCGCGTGGCGGAGAACACGCTCGTCAGCACATCGCC	AGCTATGC	3186
Db	4381	CTACATTTGACAAAGCTAACTTTCAGGCCGAAACGGGCGCGTCAGAA	CAATTCGCGATTTATGC	4440
Qy	3187	CAATCAGCAAGGCTGGACGCTGAGTAGCTGTGCGTGAAGAGCACGGCATAT	TACCGAT	3246
Db	4441	TGCGCGACAGGTTGGCGGTTACAGAGTGTGTCGCGGAAGAACAGGGTGCAT	TGCCCCAT	4500
Qy	3247	TACCTGAGGGGCAACATCGATCTGGCAACAGCAGCGCGGCGCAAGCTGCAG	CGG	3306
Db	4501	TACGTTAACGGGCGATTAATCGTCAGTTTGGCAACAGCAAC---CGCAAC	CGCTGTAGCGG	4557
Qy	3307	CCTCGCGCGCGGCTGTTTCAATCGCACACCGGTTACTCTTCGCGTCCGCGT	CGCGTGGCGCT	3366
Db	4558	ATTACGCGCGGCTGTTTTCATCGCAACACCGGCTACTCCCTACCGCTCG	CGGTGGCGCT	4617
Qy	3367	AGCGGAGTTGGTAGCAGCGCTGTGCCCCACCGATGCCCTCAGCTCAGCCAA	CATATCGA	3426
Db	4618	GGCCGATCGCTCAGCGCGCTGGATGTGTTTAACTCTTCTCTGTTTACC	CAGACGATTCG	4677
Qy	3427	ACGCTTTCCCGTTCAGCAGTGGCGGCAACAGCGATTTTTCGCTCTGCTAA	ACCGCATGCT	3486
Db	4678	TCATTTTCCCGCAACGTTTGGCAGCAACAGGGGTTTTTCGATGCTGAATCG	CAATGTT	4737







```
Db 6895 GCGATCTCGGCTGCTTCCAGTTGACCAATTCGCGGTGATATTGTGACGATGCTC 6954
Qy 5707 AAAATGGTCTGCTATCTGCGCAATCTGTGCTGATCAGCGGGATTAAGCGCGGATA 5766
Db 6955 AGGTGGCGGCTGTTATCTGCTGAAAGTGGCTGGAAGAGAAAGACTGACGAAGCGA 7014
Qy 5767 CGCTGATGTCACCGCAAAATGTCGAGCGCTGCTGCTCACTGGCAGCGGTTTATGTCGG 5826
Db 7015 ATTATGCTGCGCGCAGAAAACCGCGAGGCTTAAGCCGATGATCGCGGGGACTGGTACGG 7074
Qy 5827 AGGCGGAACCTTATATCACTCGGCGCGATCCGGTTTACCGGTTTACCGCTGCGCTCGG 5886
Db 7075 AAGCGAACCCTTATTCGATCATCAATGGCCGGTCTGGCAATATACCTTTACGCTCGG 7134
Qy 5887 CGTGGCCATCGCTACGCTCGCGCTGTTTATCGGAAATTTGGCGTCAAAAGTTTCAGACG 5946
Db 7135 CCGTGGCCATCGCGACGAGCAGGAGTGTACCGTAAATTTGGCGTGAAGTTGAACAGG 7194
Qy 5947 CCGGTGTCAGCGCTGGGATTCACGGCAGCGACACAGTAAAGGTGAAAACTGGCGCTGC 6006
Db 7195 CCGGTAAAGCGGCTGGGATCATCGCCAGTCCACGTCACCGCCGAAAAATTAACGCTTT 7254
Qy 6007 TGCTCAAGGGGACGTTTGGCGATCACTTCGCGTGTCTCGTCTGAAACGCGCTCGG 6066
Db 7255 TGCTACGCGCATCCGGTCAGGAGTACTTCCCGATGAAGACGATCCACCCCGCTCTG 7314
Qy 6067 CTGTGCTGTGGCAGCGCTCTCGTTGATTTTACGTCGCTGAGCGTGGCGAGCGTGGCTG 6126
Db 7315 CTCATCTCTGGCAGCGCC-----GATCTAGCGCATGCCCTTCTCTCAGGCTGCGCTG 7368
Qy 6127 CAGCTTATTCAGCGTGGCGGTAGAGAAACCAAACGACAGCGCTTCAACGCGCGG 6186
Db 7369 AAGTTTATAGTAACGCTGGCGGTACAGAAAAACAAAGGACACGCGAGCCCTCTTTCCCT 7428
Qy 6187 CACCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6246
Db 7429 TACAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7488
Qy 6247 ATAGCGGAACGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6306
Db 7489 GTAGCGGAACGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7548
Qy 6307 GTAGCTGATTCGCGGACCAATCACTGAGCGGCAATGCTTGGCAGCAGCAATA 6366
Db 7549 ATAAGCGGTGATGCTGCAACCAATCACTGAGCGGCAATGCTTGGTACTGCGCAAGTA 7608
Qy 6367 AATGACCAATGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6426
Db 7609 AATGAGGCAATGCGACCAATGCGGCAATGATGATGATGATGATGATGATGATGATGAT 7668
Qy 6427 ACCGCTGTGCGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6486
Db 7669 GCCTTACGCGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7728
Qy 6487 TTATGTCAGCAGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6546
Db 7729 TTATGTCAGTGCAGCAACCACTTCCATGCGGACGCGTGAACAACGATCAGGCG 7788
Qy 6547 GTTCCATAACGAGCATTTGTCGTCCA 6574
Db 7789 ATTCCAAATCCCAACATAATTTCTCAA 7816
```

## RESULT 6

```
US-10-735-442-64
; Sequence 64, Application US/10735442
; Publication NO. US20040219629A1
; GENERAL INFORMATION:
; APPLICANT: E. I. duPont de Nemours and Company, Inc.
; APPLICANT: Suh, Wonchul
; APPLICANT: Rouviere, Pierre
; APPLICANT: Cheng, Qiong
```

```
; TITLE OF INVENTION: Increasing Carotenoid Production in Bacteria Via Chromosomal
; FILE OF INVENTION: Integration
; FILE REFERENCE: CL2027 US NA
; CURRENT APPLICATION NUMBER: US/10/735.442
; CURRENT FILING DATE: 2003-12-12
; PRIOR APPLICATION NUMBER: US 60/434618
; PRIOR FILING DATE: 2002-12-19
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 64
; LENGTH: 8609
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Plasmid pPCB15
US-10-735-442-64
```

```
Query Match 37.4%; Score 2619.6; DB 18; Length 8609;
Best Local Similarity 64.7%; Pred. No. 0;
Matches 4015; Conservative 0; Mismatches 2169; Indels 24; Gaps 7;
Qy 373 CAGCTGCAGCAAAATTTTACAGCGCATCTTGAACATTTTACTGCTTCGCGACAGCAAGC 432
Db 1627 CAGTTGCTGGCTGATATCGATAGCGCTTTCATCAGTTACTGCGGTTTCAGSGTGAGCGG 1686
Qy 433 GATCGCGTGTGCGCGATGCGGAAAGCTGGCGAGGCAAGCTGGCGAGGCAAAAGTATTCGTCT 492
Db 1687 GATTTGTGGTGGCGGATGCTGTAAGCGACGCTGGCAGCGGCAACGATTTTCGTCGG 1746
Qy 493 TTATTACTGCTGCGGACGCGGATGATGGTTGCGAGCTGACGCCAAAATGGCGCTCTC 552
Db 1747 ATGCTGCTGTTTAAACAGCGCGGATCTTGGCTGTGCGATCAGTCACGGGGATTTACTG 1806
Qy 553 GATCTCGCTGTGAGTGAATGTGCAACGCGCATCGCTGATTTCTGGATGACATTCGC 612
Db 1807 GATTTAGCTGCGGTTGAATGGTGCATGCTGCTGCTGATTTCTGGATGATATGCC 1866
Qy 613 TCGATGATTAACGCGCAGATGCGTGGTGGCTTACCGTCACTCGGCAATTTGGTGAA 672
Db 1867 TGCATGACGATGCGCAGATGCTGGGGGGGTCCACCATTTCAACGAGTACGGTGAA 1926
Qy 673 AACGTGGCGATTCGCGCGCATCGCGCTGTTAGCGCGCATTTGAAGTATTCGCATT 732
Db 1927 CATGTGGGATTTGCGCGGCTGCTTTACTCAGAAAAGCGTTTGGGGTGAATGCCGAG 1986
Qy 733 GCACCGGTTTGCCTGCGCATACATAAATCTGAAGCGATTTGTGAACTCTCCGCTGCCGTC 792
Db 1987 GCTGAAGTCTGACCGCGATAGCCAAACTCGCGGCTGTCGAGCTGTCCACTGGATT 2046
Qy 793 GGCCTGCAAGGCTTAGTCAAGGGCAATTCAGGATCTGCAAGCGGCAACGAGAGCGCG 852
Db 2047 GGCATGCAAGGCTTGGTTTCAGGGCCAGTTTAAAGGACCTCTCGGAAGGCGATAAACCCCG 2106
Qy 853 AGCCCGAAGCGATCGCCATGACCAACGAACTGAAACACGCGTCTGTTTCGCGCCAG 912
Db 2107 AGCGCGATGCCATCTGCTAAACCAATCAGTTTAAACCAAGCAGCTGTTTTCGCGCTCA 2166
Qy 913 CTGCAAAATGGCGGATTTGCGCTGACGCTTTCACGCGAGGTGCGGCAAGACTTAGCTTC 972
Db 2167 ACGCAATGGGCTCCATTTGCGGCCAAACGCGTCTCGAAGCGGTGAGAACTTGCATCGT 2226
Qy 973 TTCGCCCAAGATTTGGGCGAGCGGTTTCAACTGCTGCAAGCCTCGCGCAACGCGTTCGAAA 1032
Db 2227 TTCTCGCTCGATCTCGGCGAGCGCTTTCAGTTGCTTGAATCTTACCGATGCGATGAC 2286
Qy 1033 CACACCGGTAAGATGTGACACGAGATCAGGCAATCCAGCTGCTGACAGATGCTCGGT 1092
Db 2287 GATACCGCAAGACATCAATCAGATGCAAGTAAATCAACGCTGCTCAATTTATTAGGC 2346
Qy 1093 GCTGACGCGCGGAACGCTGCGCTGCGGATCACTTCGCGCAGCGAGATGACACCTTGGC 1152
Db 2347 TCAGGCGGTGCAAGAACGCTGCGACAGCATTTGCGCCTGGCCAGTGAACACCTTCC 2406
```



Db 4558 ATTACGGCGGGCTGTTTATCCGACAAACCGGCTACTCCCTACCGCTCGCGTGGCGCT 4617  
Qy 3367 AGCGAGTGTGTAGCAGCGCTGTGTGCCACCGATGCCCTCAGCTCAGCAACATATCGA 3426  
Db 4618 GGCAGATCGTCTCAGCGCGCTGGATGTGTTTACCTCTTCTCTGTTTACACAGAGATTGC 4677  
Qy 3427 AGCTTTGGCCGTCAGCAGTGGCGGAAACAGCGATTTTTCCGTCGTCTAAACCGCATGCT 3486  
Db 4678 TCACITTTGCCAGCAACGTTGGCAGCAACAGGGGTTTTTCCGCATGCTGAATGCGATGTT 4737  
Qy 3487 GTTTTTGGCCGTTAAGCCGAGCAGCGCTGGCGCGTGATGCAACGTTTTTACCGGCTCGA 3546  
Db 4738 GTTTTTAGCCGACCGCCGAGTCAAGCTGGCGTGATGACAGGTTTCTATGGCTTACC 4797  
Qy 3547 TGC CGGGTTAAATPAGCCGTTTTTACCGCGGCAACTGCGCTGCGGATATAAACGCGGAT 3606  
Db 4798 CGAGGATTTGATGCGCGCTTTTATCGCGGAAACTCACCGTGACCGATCGGCTACGCAT 4857  
Qy 3607 TCTGTGGCGCAAGCCGCGTGGCCCATCGGTGAGCGCTCGCGCGCTGTTGAAATCTGT 3666  
Db 4858 TCTGAGCGCAAGCCGCGCTTTCCCGTTTTTCGCGGCAITTGCAAGGCAATATGACGACTCA 4917  
Qy 3667 CGAACCAAGCAAGAAATAAGAAACGCACTTATGTGATTTGGCGCAGGCTTTGGCGGCTG 3726  
Db 4918 TCGTTGAAGAGCGACTACATGAACCAACTACGTTAATTTGGTGGCGCTTTGGTGGCTG 4977  
Qy 3727 GCGCTGGCAATCGCTGCAAGCGGCGGCATACCAACCACTTACTCGAGCAGCGGAC 3786  
Db 4978 GCACTGGCAATCGTTTACAGGCGCGAGGTATTCCTGTTTTGCTGCTTGCAGCGCGGAC 5037  
Qy 3787 AAACCGGCGGACCGCGCTATGTGTTGAGGACAGTGGCTTTACCTTCATGCGGACCC 3846  
Db 5038 AAGCCGGGTGGCGGCTTATGTTTATCAGGACGAGGGCTTTACTTTGATGCGGCGCT 5097  
Qy 3847 ACGGTGATCACCGATCCAGCGCATCGAAGAGTTGTTTCAACGCTGGCAGGAAATCGCTC 3906  
Db 5098 ACGTTATCACCGATCCAGCGGATGAAGAACTGTTTGTCTTGGCCGCGTAAACAGCTT 5157  
Qy 3907 AGCGATTACGTGAGCTGATGCGGTAACGCCCTTCTATCGCTGTGCTGGGAAGATGGC 3966  
Db 5158 AAGGATTACGTGAGCTGTTGCGGTCACGCCGTTTTTATCGCTGTGCTGGAGTCCGCG 5217  
Qy 3967 AAACAGTTGATTACAGCAATATCAGCGCTGCTGGAGCAGCAGATCCCGCTTCAAT 4026  
Db 5218 AAGGTCTTCAATACGATTAACGACAGGCGGATGAAGAACTGTTTGTCTTGGCCGCGTAAACAGCTT 5277  
Qy 4027 CCGCAAGATGTAGAAGCTATCGTCAATTTCTTGCTTATTCAGTGAAGTATTTAGAGAG 4086  
Db 5278 CCGCGCATGTTGCGGGTTATCGAGCGTTTCTTGACTATTCGGTGGCGGTTTCAATGAG 5337  
Qy 4087 GGTATCTGAACCTCGGCAAGTGGCGGTTTTCTGCAAGTGGTGCATGCTGCGCGTCGG 4146  
Db 5338 GGTATCTGAAGCTCGGCACTGTGCCCTTTTATCGTTTCAAGACATGCTTCGGGCGCG 5397  
Qy 4147 CCGCAGTTGGAGCTGTCAAGCATGCGCGGCTCAGCATGTTGGCGGAAATTTATT 4206  
Db 5398 CCCAGTTGGCAAGCTGAGGCAATGCGCAGCGTTTACAGTAAAGTTGCGCGCTACATTT 5457  
Qy 4207 CAGGACGATCATCTGGTCAGGGGTTTTCTTCTCACTCATGCTGTTGGCGGTAATCT 4266  
Db 5458 GAGGATGAGCATCTTCCGACAGGGTTTTTTTCACTCGCTTCTAGTGGGGGGAATCG 5517  
Qy 4267 TTTGCAACCTCATCGATATATCTTTAATTCATGCGCTGGAGCGTGAATGGGCGGTGG 4326  
Db 5518 TTTGCAACCTCGTCCATTTATACGCTGATTACGCGGTTAGAACGGGAATGGGCGTCTG 5577  
Qy 4327 TTTCCGCGGCGGACCGCGGGCTGTGTGAGGCAATGGCGGACTGTTCAGGACTTG 4386  
Db 5578 TTTTCAACGGGGTGGAAACCGGTGCGCTGGTCAATGGGCAATGATCAAGCTGTTTTCAGGATCTG 5637  
Qy 4387 GCGCGGAGCTGTGTTACTGTAATGCCGAAGTGAGCCAGCTGGAACACGAGCGCAATCGCAT 4446  
Db 5638 GCGCGGCAAGCTGTGCTTAAACCGCGGGTCACTATATGAAGCCGTTGGGCAACAGATT 5697

Qy 4447 AGCGCGGTTCACTTAGAGGGCGGACGACGCTTCGATGCGCGCGCTGTGGCTTCCAATGCC 4506  
Db 5698 CAGGCGGTGACGTTGGAAAGACGGCAGACGCTTTGAAACCTGCGCGTGGCGTGAACGCT 5757  
Qy 4507 GACGTGTGTCATACCTACGACAAACGCTTCCGCCACCACTCCGCTGGCAATGAACGTCG 4566  
Db 5758 GATGTTGTACATACCTATCGCGATCTGTGCTCTAGCATCCCGAGCCGCTAAGCAGGG 5817  
Qy 4567 ACATCGCTGAAGCGTAAGCGCATGAGCAACTCGCTGTTTGTACTCTATTTTGGCTTGAAT 4626  
Db 5818 AAAAAGCTGCAATCCAGCGTATGATTAACCTCACTGTTTGTACTCTATTTTGGTCTCAAC 5877  
Qy 4627 CAGCGCGATGAACAGCTCGCGCACACACCGCTGTGTTTGGCCCGGCTTATCGTAGTTG 4686  
Db 5878 CATCATCAGATCAACTCGCCCATCATACCGCTGTGTTTGGGCCACGCTACCGTGAAC 5937  
Qy 4687 ATCGATGAGATTTTCAACAGCAGCAGCTGGCAGACGATTTTCTACTTTTACCTGACCGG 4746  
Db 5938 ATTTACGAAATTTTAAACCATGATGCTGTGGCTGAGGATTTTTCGCTTTATTTACACGA 5997  
Qy 4747 CCTGCGCAGCAGCATCCGTCGCTGGCACCGCCCGCTGGCAGCTTTTATGTGTTAGCG 4806  
Db 5998 CTTTGTGTACGATCCGTCCTGCTGCGACCGGAAGGTGGCGGAGCTATTTATGTCTGGCG 6057  
Qy 4807 CCGGTGCGCATCTCGGCAACCGCTGACATCGACTGGCAACAGGAAGGACCGGCTTGGCG 4866  
Db 6058 CCGTGTCCACACTTAGGCAACCGCAACCTCGACTGGCGGCTAGAAAGGACCCGACTGCGC 6117  
Qy 4867 GATCGAATTTTGTCTTATCTGCGCAGCAGCTACATGCGGGATTAAGTCAAGCAATTAGTG 4926  
Db 6118 GATCGATTTTGTGACTGCTTGTGAGCAACATTAACATGCTTGGCTTGGCAAGCAGTTGGTG 6177  
Qy 4927 ACACACAGAAATTTTACCGCGCTTTCATTTTCCGACACGCTGCATGCCCATCACGCGCTCG 4986  
Db 6178 AGCGACGTAATGTTTACCGCGCTTCGATTTCCGCGAGAGCTCAATGCTTGGCAAGTTGCG 6237  
Qy 4987 GCGTTTTGCTGAGCGCATTTTGAACGAAAGCGCTGTTTCCGCCCGCATTAACCGCGAT 5046  
Db 6238 GCGTCTCGGTTGAACCTTATCTGACCCAGAGCGCTGTTTCCGACCAACATAACCGCGAT 6297  
Qy 5047 GCGGATATCAGCAATCTCTATCTGTTGGTGGCTGCGTACGATCCAGGCGCGGCGTCC 5106  
Db 6298 AAGCACATGTAATCTTATCTGTGCGCAGGACCCATCTGCGCGGCGCATTTCC 6357  
Qy 5107 GCGGTGATCGGTTCCGCGCAAGCCACCGCAGGCTGATGCTGAGAGATCGCGCCGAATGA 5166  
Db 6358 GCGGTAATCGGCTCGCGNAGCGGACGCGAGGCTTAAATGCTGAGGAGACCTGATTTGACGA 6417  
Qy 5167 ATCGACAGCTTTTACTTGAGCAAGTAAACGAAACCAATGCGGCTGGGCTCGAAGAGTTTCG 5226  
Db 6418 AT---ACGTCAATTAATGAAATCATGCGCTCGAAACCAATGCGGCTGGCTCGAAAGCTTTG 6474  
Qy 5227 CCACCGCGCCAAAGCTGTTTGTATGACCGACCGCGCGAGCAGCTGATGCTGTATGCT 5286  
Db 6475 GCACTGCAATCGACGCTTTTCGACGCGCAAAACCGCTGCGAGCGTGTGATGCTTACGAT 6534  
Qy 5287 GGTGCTGCTACCTCGCATGATGATGATTGATGGGCAAAACGCTGGCGGAAAGCGGCAACG 5346  
Db 6535 GGTGCGGCACTGCGACGAGCTCAATGAGCATCAACACTGCGGCTTTCATGCGGACGAGC 6594  
Qy 5347 ATGCGCTCGAAGACCGCGAGGCAACGATGACGATCTGCAAAATGAAACCCCGCGCGCT 5406  
Db 6595 CCTCTTCGAGATGCTGAGCAGCGCTGACAGCGCTTGAATGAATGAACCGCTCAGGCGCT 6654  
Qy 5407 ACAGCGGCGGCAATGATGAACCGGCTTTAGGCGGTTTACGAGGTTTCCAGGAGTGGCG 5466  
Db 6655 AGCGGCTTTCGAAATGACAGGCGCGCTTTTCCGCGCTTTTCCAGGAGTGGCG 6714  
Qy 5467 ACCAGCTCGCGCAACAACTGCGGCTTGTATCTGGAAGGCTTTCGATGATGACGCA 5526  
Db 6715 ATGATATCGCTCCCGCTACGCGTTTCGACCATCTGGAAGGTTTTGCGATGATGCGCG 6774

[illegible]

RESULT 7  
 US-10-808-807-18  
 ; Sequence 18, Application US/1080807  
 ; Publication No. US20040253663A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: E.I. du Pont de Nemours and Co., Inc.  
 ; APPLICANT: Cheng, Qiong  
 ; APPLICANT: Tao, Luan  
 ; APPLICANT: Sedkova, Natalia  
 ; TITLE OF INVENTION: GENES ENCODING CAROTENOID COMPOUNDS  
 ; FILE REFERENCE: CL2365 US NA  
 ; CURRENT APPLICATION NUMBER: US/10/808,807  
 ; CURRENT FILING DATE: 2004-03-24  
 ; PRIOR APPLICATION NUMBER: US 60/477,874  
 ; PRIOR FILING DATE: 2003-06-12  
 ; PRIOR APPLICATION NUMBER: US 60/527,083  
 ; PRIOR FILING DATE: 2003-12-03  
 ; NUMBER OF SEQ ID NOS: 20  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 18  
 ; LENGTH: 8814  
 ; TYPE: DNA  
 ; ORGANISM: Pantoea agglomerans strain DC404  
 US-10-808-807-18  
  
 Query Match 30.6%; Score 2138.6; DB 18; Length 8814;  
 Best Local Similarity 60.7%; Pred. No. 0;  
 Matches 3811; Conservative 0; Mismatches 2279; Indels 185; Gaps 12;

Qy	309	ACCATGATAGCCATTATGACCCATGTCGACACACAGACACATCAGACAAAGCGAACT	368
Db	2384	ACGCCGGTACCAACCATGACAAGACCCCTTTGAAACACATCCCGGTCAAGACGGGAACT	2443
Qy	369	CCTTTCACTGCAGCAAAATTTTACAGGCGCAATCTTGAAACATTTACTGCTGCCGGACAGCA	428
Db	2444	GCATGAGCTGCACGCTGCCCTGCAACCGTCGCCCTGGATGAACCTGCTGCCCTTTGGCGATGA	2503
Qy	429	AAGCGATCGCGTGGTCCGCGATGGTGCAGAAACGCTGGCGCAGGGCAACAGTATTCG	488
Db	2504	GCGGGATCGGGTCAGCAGCGCAATGCGGAAGCGGTACTGCGACCCGGGGAACGCATTCG	2563
Qy	489	TCCTTTTATTACTGCTGTCGACGCGCGCATATGCGGTGTCAGCTGACGCAAAATGGCGT	548
Db	2564	CCGCGTCTCTGTATCTTCGCGCGCCGCGACCTCGGCTCGCATCGCAACACCCCGGCT	2623
Qy	549	TCTCGATCTCGCTGTGACGTGAATGTCACGGGCATCGCTGATTCCTGGATGACAT	608
Db	2624	GCTGGATATGGCCCTGTCGGTGAATAATGGTGCAGCCCTCTGCTGATCTCTGACGATAT	2683
Qy	609	TCCCTTCGATGATAACCGCGCAGATGCGTCTGTTGTCGCTTACCGTGCATCGCGAATTTGG	668
Db	2684	TCCCTTCGATGATAACCGCGGCTCCGCGCGGTCCGCTTACCATTCATCGCCAGTATGG	2743
Qy	669	TGAAAACTGGCGATTCCTCGCGGCATCGCGCTGCTTAGCGCGCATTCGATGATTTC	728
Db	2744	TGAAGAGTGGCAATTCGCTCGCGTAGCGTGTCTCAGCAGCGCCCTTTGGCGTGTGTT	2803
Qy	729	CATTGCAACCGGTTTGGCTGCGCATACATAAACTGGAAGCATTTGCTGAACTCTCCGCTGC	788
Db	2804	CGCGCGCAGGATTGTCTCCGAGTGCCGACGCGGTGGCGAGCTGCGATGCG	2863
Qy	789	CGTGGCCCTGAGGGCTTAGTGGAAGGGCAATTCACAGATCTGCACACGCGCAGCAGAG	848
Db	2864	GCTCGGTATCCCAAGGCTTGGTGTGAGGGTCAGTATPAGGATCTGCGTGAAGGCACCGCCC	2923
Qy	849	CCGACGCCCGGAAGCGATTCGCATATGACCAACGAACCTGAAAAACAGCGTGTGTTTCGCG	908
Db	2924	CGCAGCGCCGAGGAGATGCGCACCAACGAATGAAACACAGCGTGTGTTTGGTGC	2983
Qy	909	CACGTGCAAAATGGCGGATTGGCGTTCAGCTTCAACCGAGTGTGGGCAAGACTTAG	968
Db	2984	CACGTGCAAAATCGCGCCCTGGCGGAGGCGCTCGCGCGCGCGCGCGGAAATTCGC	3043

QY 969 CTTCCTTCGCCAGGATTTGGGCCAGGCGTTTCACTGCTCGAGACCTCGCCGAGCGGTG 1028  
DB 3044 CTGCTTTGGCGAGGATTTAGGCGAGCGGCTTCAGCTGCTGGACGATCTGGCGGACGGCCA 3103  
QY 1029 CAAACACACCGGTAAAGATGTGCACAGGATCAGGGCAATCCACGCTGCTACAGATGCT 1088  
DB 3104 TGCCGGGACCGGCAAGACATCAATAAGACGGGGTAAGTCCACGCTGGTGGCGATGCT 3163  
QY 1089 CGGTGCTGACGGCGCGGAAACGCTGCTGCGCGGATCACTTCGCGAGCGCAGATGCACACT 1148  
DB 3164 CGGACGAGCGGCTGCGGAGCGGCTCGACACCCATCTGCGCGCGGAGAGCGCCATTT 3223  
QY 1149 TGCGCTGCGCTGCCATCGCGGACATCGCCACTCGCCAAATATATGACGCGCTGTTAATCA 1208  
DB 3224 TTACGCGCTCGGGAAGAAACAGGCGCAAGCGCTTTATGACGCGCTGGTTTCAAA 3283  
QY 1209 ACAGCTAGCGATATCAACTGAGCGCGGCTCAGCGGTGGGCGACATTCGCGGTGATCGCG 1268  
DB 3284 ACAGCTGGCGGCTTTAGCTGAGCAACGATACACCCCGGTAAATATTTGTGAGATCACA 3343  
QY 1269 CGCGCGCTCTACAGCCACTTTTACGCGTTTCAGGCGTTAGCACAAACGCTGCTGGCGCG 1328  
DB 3344 T-----GA 3346  
QY 1329 GGCCATCGCATCACATTCATCCAGCAAGCGGATGCCGCGACTTTGTCTTAGCGACGACGC 1388  
DB 3347 AGGACGCGATCTGGTTTACGCGTAAATAATGACACCTGATATCTGCTGCAACCTTGACC 3406  
QY 1389 ATCGATTTTGTTCGCTGCGGCAACAGACGATCTCGCGTTTCGCTGGCGCGCGGTG 1448  
DB 3407 GGGCGATGATACCATTTGCAACGGATTTGACGCTGGGTTTGAACACTGC----- 3459  
QY 1449 CATCGCTGCGCTCGCGGCGGCTGTGCTGTTTTCGCGTGTATCGACGATCTCGCGTCC 1508  
DB 3460 -----GCCCTCCGGAGCTGATCTCGACGGTATCGATCTCTCCACC 3501  
QY 1509 TGCACCGATATGCTGTGCGCGNACTGCTGCGGTACTGAAGCATTTGAACATCGATGCG 1568  
DB 3502 ACCCTGTTTTCCGCGCGCTGAAAGCCCGCGTCTGATCAGCTCCATGACCGCGCGCGG 3561  
QY 1569 GTGATCGCGACGAAATGAAGCGGCGGATGCTGCTGTAAGCGCTGCACTGCGCG 1628  
DB 3562 GCGCGCGCCAGAGACATTAACGCTCATCTGGCCAGCGCGGCAACCTTTGGGCTGGCG 3621  
QY 1629 TTGTTTTCGCTGCGCTTCGCGCTTCGCGGTCAATCGTAAGCGGGAATTCGCTTGGG 1688  
DB 3622 ATGGGCGTTCGTTCCAGCGGCTGGCGCTGGAGGACGGGCGGACGCGGCTGGATGCC 3681  
QY 1689 ATGCCCTTCGTTTTCACAGGATGACAAAGCGCTGAAACGTTTTCAGGCCAGCAGCAT 1748  
DB 3682 CAGCTACGCCATATCGCCCGGACGTGCGGCTGCTG-----GCT 3720  
QY 1749 ATCTATGATCGCATATGCTGCTACGCGGCGGTGATCTCTCAACACGCGGCGGCTTT 1808  
DB 3721 AACCTTGGCGCGCGAGATCCCGGTGCGAGGGCTTGACTACGCCGCGCGCGGTG 3780  
QY 1809 AATTGACGAGGCGCGGATTAATCATGATGCTGCTGCGCGCTGGCGACAAATCAGCCAG 1868  
DB 3781 GACATGATCGACGCGAGCGGTTAATTGTGATCTGAAACCGGCTG-----CAG 3828  
QY 1869 ATGGTCCGGCTTTGATTTTTCAGCTCAGCAATCGCCGCTGCTGATACGCGGTGGGG 1928  
DB 3829 GAGGCGCTCCAGGCGCGCGGATCGCGATCGGCGGATCTCTCAACGCCATTTGCGCAG 3888  
QY 1929 CCACCTCGGCGCGGTTTCTCTGCGCGCTCCATGCGCCCTGGCGAGCGCTGCGTCAG 1988  
DB 3889 CTGGTGGCGAATCGCGGTACCGGTGCGGTGTTAAAGAG--GTGGGCGCGGATCTCCC 3946  
QY 1989 CCGGTGTTTATGCTTCGCTGCGGTAGCTGCAAGGCCATGCTGCTCGCGCTGTTTCTGCA 2048  
DB 3947 CGGACGTTGCTGCGCAGTGGCGGAGCTCGGCGTGGCGATGATCGACATTG----- 3997  
QY 2049 CTGGCGAGCGGTGCCCGCAGCTGCGGCTATCGCTGCTGATCGCCCATTTGTGGGGATTA 2108

DB 3998 CCGCGCGCGGCGGAAACACAGCTGGGCGGCGGTGGAAGCTGAAACGCGCCCCG----- 4047  
QY 2109 AACGCGCGAACACAGCATCAGCTGGAGCTCGCTGGCGCGGCTGGGTGACGGAATTTGCTC 2168  
DB 4048 ---ACCCCGAGGCGGAAATGTGGGATGGCTTTGCCGACTGGGCAATCTCTACTGCC 4104  
QY 2169 GATCAGCGGCGACCCCTACAGCACGCGGAGCTGTTTATCATTCACTCAGCGGGTTAAACAGC 2228  
DB 4105 GATGGCTGCGTTCGCGTCCATCTTGGCTGCCGTGATATCCCGCTTATCGCCTCCGCGGC 4164  
QY 2229 GCGCTGGAGACTCGAATGCGGTACGCGGATGCTGGCGCTGCGGATGCTTTTGTATCAG 2288  
DB 4165 ATGCGCAACGCGATTTGACGCGACAA-----AAGCCATCGCGCTGGGTGAGATCTG 4215  
QY 2289 CCGCGCTGGCGGCGCGCATTTGAGTGGCATGAGCTGTGTGCGCGGCGCATCAAGCTTTAGC 2348  
DB 4216 GTGGGCGAGCGCGCGGTGCTGGCGCATGCAACGCCCTCGCGGACGCGGCAATTGCC 4275  
QY 2349 CGTGTTCATCAACTGGAGCAGCATCTGCAACAGCTGCTGACCGACGATGTTACGCGCTA 2408  
DB 4276 CATTTCCGCACTTATACGCACTGC-----GGATCGCTGTTTCTGTACCGGCAG 4328  
QY 2409 CGGATGTCAGCGATTTACGCGCAGCTGCGCGGAGCGGTTGCCAGGCTGCGCGCGAC 2468  
DB 4329 TGCAAACTGCAAGCGGTTCGACACGCCACGCTGCTTCGCGTCAACGGCGGCGCATCCCT 4388  
QY 2469 ATGCTCGAGCAGCGCTGTCGCGAGCAAGCTGCTGCTGGCGGAGCGACCTGATCGCA 2528  
DB 4389 GTGACGAGTACGCTGCTTATACCGGGAGCGGT-----ATGA 4427  
QY 2529 CGCAATAAGATGTGATTTTGGTGGTCTGGAATGCGGAAATGGCTTGTATGCGCTGCGTC 2588  
DB 4428 AAAAAATGGGATCTGATTTCTGGTGGCGGGGCTGGCCAAACGGGCTTATCGCTGCGCAC 4487  
QY 2589 TCGGCTCAATTGCGAGCACCACTGAATGCTGCTGCTGGAGAGCGATGCGATCCGCGCAG 2648  
DB 4488 TAAAGAGCGCTCATCCGACGCTTGTGTTAATGCTGGAGTGGCGGAGCGGCCCGCGG 4547  
QY 2649 GCAATCATACCTGCTGTTTTCATACAGCGATCTCAGCGCCGAAACAACTTCGCTGCTGC 2708  
DB 4548 GAAACCAACACTGCTGCTTTCACCAACAGCATATCAGCGCCAGCCAGCAGCGCTGCTG 4607  
QY 2709 AACCGCTGATTAACGCTGGTGGTTCAGGTTATCAGGTGGTGTTCCTGCGCTGCGCGCA 2768  
DB 4608 CGCGCTGCTGGCCCATCGCTGGGACGGGTACAGCTCCACTTTCGGAACGTGCTCGCA 4667  
QY 2769 ATCTGAGCGGGATTTATTTCCATCGCATCAGGCGATTTTCCCGCCATCTTTACGCG 2828  
DB 4668 CCTCGATGACGGCTACTGACCATCACTCCACGCTTTTCCCAAGCGATGCGCGGGC 4727  
QY 2829 CGATGGGTGACCATCTGTGGACAAACACAGCGGTACACAGGTAAACCCAGCGAGGTGA 2888  
DB 4728 TGATGAAGAGATTTTGTGACAAACGTCAGCTGCTCAGGGGTAGCGGCGAGGAAGTAA 4787  
QY 2889 CGCTGGCGGATGCGCTGAACTTGTGCGCAAGTGGTGAATGATGCTCGCGGCTGCGAGC 2948  
DB 4788 CCTCAGGACGAGACGCTTTTACCGCGGCGGTGATGATGCGCGGCTATCAGC 4847  
QY 2949 CGAGCCCATCTGCACTGGGTTATCAGGTGCTTTCTTGGAACAGTGGAGCTGCGCGC 3008  
DB 4848 CCTCGCGCATCTCAGCATTTGGCTATCAGGCGTTTATCGGCGAGGATGGCACTGACCG 4907  
QY 3009 AGCCGCGAGCGCTGCGAGCAGCGATCTGATGGATGCCACCTGATCAGCAAGCGGGTT 3068  
DB 4908 CGCCCAACGGGTTAACCGCGCCGATCTGATGATGATGCGCGCTCGCCAGGCAACGGCT 4967  
QY 3069 ATCGTTTGTCTACAGCTGCGCTCAGCGCGATCGGCTATTGATTGAAGATACCATTT 3128  
DB 4968 ACCGCTTTGCTATATACCTGCGGCTCAGCGCGACACCTGCTTATCGAAGACACGCACT 5027  
QY 3129 ACGTTAACCGCGCTGGCGGAGAACACCGCTGCTCAGCACATCTGCGGCTATGCGCA 3188



Db 5028 ACATTGACGGCCCGGACGCTGCGACGCCGATTGACGCCCGCGCGGATTGCCGATTACGCC 5087  
Qy 3189 ATCAGCAAGGCTGGACGCTGAGTACGCTGCTGCGTGAAGAGCACCGCATATTACCGATT 3248  
Db 5088 GCCAGAGGCTGGACGCTGCGCGCTGGTGGTGGAGAACAGGGGGCGCTGCCGATCA 5147  
Qy 3249 CCTGAGCGGCAACATCGATCGATTCTGGCAACAGCGGGCCCAAGGTGCGAGCGCC 3308  
Db 5148 CCTGTCCGCGATCCGGCGCGCTTCTGGCACAGTCCATCATCAGCCGCTCAGCGGCC 5207  
Qy 3309 TCGCGCGCGGCTGTTTATGACCAACACCGGTTACTCTTGGCGCTCCGCGTGGCGCTAG 3368  
Db 5208 TGGCGCCGCTGTTTTCATGACCAACCGGCTATTGCTGCGCTGGGGTGGGCTGG 5267  
Qy 3369 CGGAGTTGTAGCAGCGCTGTTGCCAACGATGCCCTCAGCTCAGCGCAACATATCGAAC 3428  
Db 5268 CGGACCGCATTCGCAACGCGCGGACTGCATCAGGGCGCGCTCTATCAGCTGATCGCG 5327  
Qy 3429 GCTTTCGCTGACGAGTGGCGGACAGCGATTTTCCGCTGCTGCTAAACCGCATGCTGT 3488  
Db 5328 ATTTCCGCGCGCGCTGCTGGCAGACACAGCGTTTTCGCGCTGCTTAAACCGCATGCTTT 5387  
Qy 3489 TTTTGGCCGCTAAGCCGCGAGCGCTGGCGCTGATGCAACGTTTTTACCGGCTCGATG 3548  
Db 5388 TCTTGGCGGCAACACCGACCGAGCGCTGGCGGCTGATGACGCGTTTTTACCGATGACG 5447  
Qy 3549 CCGGTTAATTAGCCGCTTTTACCGCGGCACTGGCGCTGGCGATAAACCGCGATTC 3608  
Db 5448 AGCAGCTGATCGCCCGCTTTTATGCGCGCGAGCTTCGCTCCCGCGACCGCGCGCGTGC 5507  
Qy 3609 TGTGCGGCAAGCCGCGCTGCCCATCGGTGAAGCGCTGGCGCGCTGTT-----GAATT 3662  
Db 5508 TGTGCGCAACCGCGCTGCGGATTTGTGGGGGATCAAGCCCTGCTCCACACTCAATT 5567  
Qy 3663 CTGTGCAACACAGGGAAGAAATGAAACGCACTTATGTGATGGCGCAGGCTTTGGCGG 3722  
Db 5568 CTTCCTCGAGCGCCATCAATAATGAAACAAACCAATTGTAATTGGCGCGCGGTTCCGCGG 5627  
Qy 3723 CTTGGCGCTGGCATTCGCTGCAAGCGGGGCATACCAACCACTTACTCGAGCAGCG 3782  
Db 5628 ACTGGCGCTGGCATTCGCTCCAGGCGGGGCATTCCTACACGCTGCTGGAGCGCG 5687  
Qy 3783 CGCAAAACCGGGGCGGACGCGCTATGTGTTGAGGACAGTGGCTTTACCTTCGATGCCGG 3842  
Db 5688 CGCAAAACCGGGCGCGCGCTATGTCTAGGAAGATCGCGGCTTTACCTTTGATCGCGG 5747  
Qy 3843 ACCACCGTGATACCGATCCAGCGCATCGAAGATTGTTCAAGCTGCGAGGAAATC 3902  
Db 5748 TCCACCGTATCACCGATCCCTCCGCAATTGAGGAGCTGTTCAACCCTCGCGGAAACG 5807  
Qy 3903 GCTCAGCGATTACGTCGAGCTGATGCCGTAACGCGCTTCTATCGCTGTGTCTGGGAAGA 3962  
Db 5808 GCTGAGGACCTAGCTTGAAGCTGATGCGGTTGACGCGCTTCTATCGCTGTGTCTGGGAAGA 5867  
Qy 3963 TGGCAAAACGCTTGAATTACGCAATATACGCGCTGCTGGAGCAGATCCGACGTT 4022  
Db 5868 CGGCAAGGTTTTCGACTACGCAACGATCAGCGGGCGCTTGAATCGCAGATCCCGCGTT 5927  
Qy 4023 CAATCCGCAAGATGAGAGCTATCGTCAATTCTTTCGCTATTTCAGCTCAAGTATTTAG 4082  
Db 5928 TAAACCGCAAGACGATGGCGGGCTATACCGCTTCTCGACTCTCCGCGGCTGTTTCG 5987  
Qy 4083 AGAGGTTATCTGAAATCCGCGACGCTGCGGTTTCTGAGGTGCGTGACATGTGCGGCT 4142  
Db 5988 CGAAGGCTATCTGAAGCTCGGCGGGTTCGCTTCTCTGCTTTCGCGACATGCTGGCGC 6047  
Qy 4143 CGCGCGGATTTGGGACGCTGGAAGATGCGCGAGGCTCTACAGCATGCTGGCGAAATT 4202  
Db 6048 CGGTCCTCAACTGGCGGGCTGCAAGGATGCGCGAGCGGTGACGACAAAGTTCGGGCTTA 6107  
Qy 4203 TATTACGAGCATCATCTGCGCTCAGGGGTTTTTCTTCTCACTCAATTGCTGGTGGCGGTAA 4262  
Db 6108 CTGGAGACGAGCACCTGCGCGCAGGCAATTTTCGTTTCACCTCGCTGCTGGTGGCGGCAA 6167

Qy 4263 TCCCTTTGCAACCTCATCGATCTATACCTTAATTATCATGCGCTGAGAGCTGAATGGGCGCT 4322  
Db 6168 CCCGCTTCCACGCTCTTCTATTATACCTGATCAGCGCTTGGAGCGGGAATGGGCGCT 6227  
Qy 4323 GTGGTTTTCGCGCGCGGCGCACCGCGCGCTGTGTGCAAGGCGATGCGCGCATGTTCGAGGA 4382  
Db 6228 CTGGTTCCCGCGCGCGGCGCACCGCTGCGTGTGTGCAAGGCGATGCTGAAAGCTGTTTCAGGA 6287  
Qy 4383 CTTGGGGGCGAGCTGTACTGAAATGCCGNAAGTGAGCGAGCTGGAACACGACGCGCAATCG 4442  
Db 6288 TCTTGGCGGCAACCTCACCTTAAACGCTCAGGTTGAGCGGCTGAGAGCGGTGACAAATCA 6347  
Qy 4443 CATTAGCGCGCTTCAGTTAGAGGCGGAGCACGACGCTTCGATGCGCGCGCTGTGCGCTCCAA 4502  
Db 6348 GGTGAGCGCGTGCATCTGCTTAAACGCGGCGACGCGCTGGAGGCTGCGCGGCTGCGCTCGAA 6407  
Qy 4503 TGGCGAGCGTGGTGATACCTACGACAAACCTGCTTCGCGCACCATCCGCTGCGCAATGAACG 4562  
Db 6408 CGCGAGCTGTGTAATACCTATGCGCGACTGCTCGGCCATCACCCGCGACGCGCGCGCTAC 6467  
Qy 4563 TGGGACATCCGCTGAAGCGTAAGCGCATGAGCAACTCGCTGTTTGTACTCTATTATTGGCCT 4622  
Db 6468 GGCRAAAAGCTGAAACGCAAGCGCATGAGCAACTCGCTGTTGCTGCTCTATTATTGGCCT 6527  
Qy 4623 GAATCAGCGCGCATGAAACAGCTTCGCGCACCAACCGCTGTTTGGCCGCGTTATCGTGA 4682  
Db 6528 GGATCACCATCACACCAGCTGGCGCACCATACCGTCTGCTTTGGCCGCGGTTATAAGC 6587  
Qy 4683 GTTGATCGATGAGATTTTCAACAGCAGCAGCTGGCAGACGATTTTTCACCTTACCTGCA 4742  
Db 6588 GCTAATCGATGAAATTTTTCAGCGCGCGACACCCCTGTCGGAAGATTTTTCGCTCTATCTGCA 6647  
Qy 4743 CGCGCTTCGACAGCGATCCGCTCGCTGGCACCCCGCGCTGGCGGAGCTTTTATGTGT 4802  
Db 6648 TGGCGCTTGGCTAACCGACCGCTGCTGCGCGCGCGGCTGGCGGAGCTACTATGTGCT 6707  
Qy 4803 AGCGCGGCTCGCATCTTGGCACCGCTGACATGCTGACATGCGCACAGAGAGCGCGCTT 4862  
Db 6708 CGCGCGGCTGGCGCACCTCGGTAAACGCGCGCTGCTGAGAGCTGGAAGGCGCGCTCT 6767  
Qy 4863 CGCGGATCGAATTTTGTCTATCTGGAGCAGCATACATGCGCGGATTTTTCGTCAGCAATT 4922  
Db 6768 GCGGATCGCATTTTGTATCTGAGCGCGCTATGCGCGGCTGCGCTCCAGCT 6827  
Qy 4923 AGTGACACAGAAATTTTACCGCGTTTGAATTTTGGGACACGCTGATGCCCATCAGG 4982  
Db 6828 GGTGACGCAACCGCATTTTTCAGCGGGAAGATTTTTCGCGATACGCTCGATGCTGGCAGG 6887  
Qy 4983 CTCGCGGTTTTCGCTGGAGCGGATTTTTCAGCAAGAGCGCTGTTTCGCGCGCATAAACG 5042  
Db 6888 GTTCAGCGTTTTCATCTGGAGCGGATCTCTACCGAGCGCTGTTTCGCGCGCACACCG 5097  
Qy 5043 CGATCCCGATATCAGCAATCTCTATCTGGTGGGTGCGGTTACGATCAGCATCAGCGCGGCT 5102  
Db 6948 CGACAGCGTGGTTGATAACCTCTACTGTTGGCGCGGAACGATCCCGCGCTGGCGT 7007  
Qy 5103 GCGCGGCTGATCGGTCAGGCGCACCGCGAGCTGATGCTGGAGGATCGCGCGA 5162  
Db 7008 GCGCGGCTGATCGGATCGGATCGGCAAGGCAACCGGCTTAATGTTTAAAGGATTTAGCGTA 7067  
Qy 5163 ATGATCGACAGCTTTTACTTTGAGCAAGTAAGCAAAACCATGCGGCTGGGCTCGAAGAGT 5222  
Db 7068 ATG----TCCAGCGCGCTTCTGAAACACCGCAGCGCCACCATGACCGCGGTTCTAAAAGT 7124  
Qy 5223 TTTGCGCACCGCGCGCAAGCTGTTGATGCAACGAGCGCGCGGAGCACGCTGATGCTGAT 5282  
Db 7125 TTTGCGCACCGCTCAAAGCTTTTGACAAACGCAACCGCGCGCAGCGCTGATGCTCTAT 7184  
Qy 5283 GGTGGTGTCTGCTACATGCGATGATGATGAGGCAAAACGCTGGCGGAGCGCGGACG 5342  
Db 7185 ACCTGGTGCCTACTGCGACGATGTTATTCGAGCAGAGGTTGGTGGGTTTGTCTGCCCG 7244







8447 GAGCGTCTACCAAGTCGGTTTCGGCTTTATTGAGGATGAGCATCTGCGGACGGCCTTCTC 8506  
4236 CTTTCCACTCATCTGCTGGTGGCGGTAAATCTTTTGCAACGTTCATCGATCTATACCTTAAAT 4295  
8507 GTTTCACCTCCCTGCTGGTAGGGCGGCAACCCCTTCACCACTCGTCCATCTACACCTGAT 8566  
4296 TCATGGCTGGAGCGTGAATGGGCGGTGTTGTTTCGCGCGGGGACCGGCGCGCTGGT 4355  
8567 CCAACGCTTTGAGCGGGAGTGGGGGTCTGGTTCCCTGAGGGCGGACCGGGGCGCTGGT 8626  
4356 GCAGGCGATGGCGCGACTGTTCCAGGACTTGGGCGGCGAGCTGTTACTGAATGCCGAAGT 4415  
8627 GAAAGGATGGTGAAGCTGTTTACCGATCTGGGCGGGGAGATCGAACTCBAAGCCGGGT 8686  
4416 GAGCAGCTGGAAAAACAGCGGCAATCGCAATTAGCGCGCTTCAGTTAGAGGGCGGACGACG 4475  
8687 CGAAGAGCTGGTGGCGCGATAAACCGCGTAAAGCCAGGTCCGCGTGGCGGATGGTCCGAT 8746  
4476 CTTGATGCGCGCGCTGGCTTCCATCGCACTGCTGAGCGTGGTGCATACCTACGACAACTGCT 4535  
8747 CTTTGACACCGACCGCTGAGCTTCGAAACGCTGACGTGGTGGTGAACACCTATANAAGCTGCT 8806  
4536 TCGCACCACTCCGCTGGCAATGAAACGTCGACATCGCTGAAAGCGTAAGCGCATGAGCAA 4595  
8807 CGGCGACCAATCGGTGGGCGAGAGCGGCGGCGAGCGCTGGAGCGCAGAGCATGAGCAA 8866  
4596 CTCGCTGTTTGTACTCTATTTTGGCTGAATCAGCGCGATGAACAGCTCCGCGACACAC 4655  
8867 CTCGCTGTTTGTCTCTACTTCGCGCTGAACACGAGCTCATTTCCAGCTGGCGCACCATAC 8926  
4656 CCGTCTGTTTGGCGCGCTTTCGTGAGTTGATCGATGAGATTTTCAACAGCAGCGAGT 4715  
8927 CATCTGTTTGTGCTCCCGCTACCGGGAGCTGATCGACGAGATCTTTTACCGGCGAGCGCT 8986  
4716 GGCAGACGATTTTCACTTTTACTGCAACGCGCTTCGACAGCGATCGCTGCGACAC 4775  
8987 GCGGATGACTTCTCGCTCTACTGCACTCGCCCTGCGATCGGATCCCTCGCTCGCGC 9046  
4776 GCCCGCTCGGACGCTTTATGTTAGCGCGGTGCGGATCTCGGACCGCTGACAT 4835  
9047 TCCCGCTCGCGAGCTTCTAGTGTGCGCGCGGTGCGGATCTTTGGCAAACGCGCGCT 9106  
4836 CCACTGGCAACAGGAAGACCGGCTTGGCGATCGAATTTTGTCTTATCTGAGGACGA 4895  
9107 GGAATGGGCGAGGAGGGGCGGAAGCTGCGCGACCGCATCTTTGACTACCTTTGAAAGCG 9166  
4896 CTACATGCGCGATTTACGTGAGCAATTTAGTGACACAGAAATGTTTACGCGCTTTGATTT 4955  
9167 CTATATGCGCGCTGCTGAGCGAGCTGTGTGATCCAGCGGATCTTTTACCGCGAGCTT 9226  
4956 TCGGACACGCTGCATGCCCCATCAACGCTCGGCTTTTCGCTGGAGCGGATTTTGACGA 5015  
9227 CCAAGACACGCTGGATGCGCATCTGGGATCGGCTTCTCCATCGAGCGCGCTGTGACCCA 9286  
5016 AAGCGCTGTTCCGCGCGCATAAACCGCATGCGGATATCAGCAATCTCTATCTGGTGG 5075  
9287 AAGCGCTGTTCCGCGCGCATAAACCGCATGCGGATATGCGGATCTCTATCTGGTGG 9346  
5076 TCGCGGTACGATTCAGGCGCGGCGGTGCGCGGTGATCGGTTGCGGCAAGCGCACCGC 5135  
9347 CGCAGTACTCACTTGGGCGGGCAATCTCTGCGGTAGTGCGCTCGGCGAAGACCAACCGC 9406  
5136 CAGGCTGATGCTGGAGGATCGCGCGGAATGAATCGACAGCCTTTTACTTGAGCAAGTAAGC 5195  
9407 CAGCCTGATGATTGAGGATC---TGCAATGAGCCAAACCGCGCTGCTTGACCAAGCCACG 9463  
5196 CAAACCATGGCGGTGGGCTCGAAGAGTTTGCACACCGCGCGCAAGCTGTTGATGACCGC 5255  
9464 CAGACCATGGCCAAACGGCTCGAAAGTTTGGCACCGCTGCGAAGCTGTTTGAACCGCGCC 9523  
5256 ACGCGCGGACGACGCTGATGTTATGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5315  
9524 ACCGCGGTAGCTGTGATGCTCTTACACTGCTGCGCGCACCTGCGATGACGCTATTGAC 9583

5316 GGGCAAAACGCTGGCGGAAGCGCGCACGACGATGCGCTGCAAGACGCGCAGGCACTATG 5375  
9584 GACCAGACCCAGCGCTTCCACGAGGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 9643  
5376 CAGCATCTGCAAAATTTGAAACCGCGCGCTACAGCGGCGCGACATGCAATGCAACCGGCG 5435  
9644 GCCCGGCTGCGCACGCTGACCTGCGGCGGTTTGAAGGGGCGGAGATGAGGATTCGGCC 9703  
5436 TTTAGGCGGCTTTTACAGAAAGTGGCGATCAATTTACACGAGCTGCCGCAACAACTGGCGTTGAT 5495  
9704 TTCGCTGCTTTTACAGAGTGGCGCTGACCCACGCTATTTACGCCCGCATGGCGCTCGAT 9763  
5496 CATCTGGAAGGCTTCGCTATGAGATGCAACGAAACGAAATTAACGAGAGCTTTCGATGACAG 5555  
9764 CACCTCGACGGCTTTGCGATGAGCTGGCTCAGACCGCTATGTCACCTTTGAGGATACG 9823  
5556 CTGCGTACTGCTATCAGCTGCGGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 5615  
9824 CTGCGCTACTGCTATCAGCTGCGGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 9883  
5616 GTGCGCGACGAGCGGTGCTCGATCAGCGCTCGGATTTTAGGACTGGCGTTCCAGCTCACT 5675  
9884 GTGCGGATGAGCGGTGCTGATCGCGCTCGGATCTGGGCTGGCTTCCAGCTGACG 9943  
5676 AACATTGCGCGCAATTTGTAGAAATGCGCGAAATTTGCTGCTGCTATCTGCGCGCAATCC 5735  
9944 AATATCGCGCGGATTTATTGACGATGCGGCTATTGACCGCTGCTATCTGCGCGCGAG 10003  
5736 TGGCTCGATCAGCGGATTAACGCGCGATACGCTGACGCGCAACGCTGCGCAACATGTCGAG 5795  
10004 TGGCTGCGAGATGCGCGGCTGACCGCGGAGAACTATGCGCGCGGAGAAATTCGGGCGCG 10063  
5796 CTCGCGCTCACTGGCAGCGCTTTAGTGGCGAGGCGGAAACCTTATCTCACTCGCGCGCA 5855  
10064 CTGCGCGGCTGGCGAGCGGCTTATGATGCGCGAGAGCGCTACTACTCTCTCCAG 10123  
5856 TCGCGTTTACCGGCTTTACGCTGCGCTGCGGCGGCGCATCGCTACGCGCTCGCGCGTT 5915  
10124 GCGGCGCTACAGATCTGCGCGCGCTGCGCTGCGGCGGATCGCCACCGCGCGAGCGTC 10183  
5916 TATCGGAAATTTGGCGTCAAAAGTTTACGACGCGCGGTGTCAGCGCTGGGATTTACGCGCAG 5975  
10184 TACCGGAGATCGGTTAAGGTAAAGCGCGGAGGAGCGCGCTGGGATTCGCCCGCAG 10243  
5976 CCGACAGTAAAGTGAATACTGGCGCTGCTGCTGTAAGAGGCGGAGGTTTGGCGATCACT 6035  
10244 CACACGAAAGTTGAAATAATTTGCCATGCTGATGGCGCGCACCGGCGAGGTTATTCCG 10303  
6036 TCGCGTGTCTGCTCTGAACCGGCTCGGCTGGTCTGTGGGAGGCGCTCTCGTTGATTT 6095  
10304 GCGAAGACGACGAGGCTGACGCGCGCTCGCGCGGCTTTTGGCAGCGTCCCGTTTAGGCG 10363  
6096 TACGTCGCTGAGCTGGCGCGGCTGGCTGGCTGAGCTTATTCAGCGGTGGCGGTAGAGGA 6155  
10364 GCGCGCATGATGATTTACGCGAGATCGCTGTAGTGGCAGGCTTTCGGGCGTAAATAA 10423  
6156 AACCAAAACGACACGACGCTTTCACGCGCGCGCACCGCATGATGATGCGGTGGCGCATGT 6215  
10424 AACCGAAGGAGACGAGCGCTTCCGCGCGCGCACCGCGTGGTGGAGCGGTGGCGACGT 10483  
6216 ATAAGCGCTTAAGATGAGCTTTTGGCGGATATAGCGGAAACGCGCGAGCGTTGATGACCA 6275  
10484 AGAGCGCTTACGTTAGTAGCGCGCGGATCCAGTGAAGGCGCGAGCGCTGATGACCA 10543  
6276 GCGCATCTGACCATGAGTAGGCGCGCGTACGTCGTCATTCGGGCAACCAATCCACT 6335  
10544 GACCGTGTGTCACAGGAGTAGAGGAGGAGTAGAGGAGGAGGAGGAGGAGGAGGAGGAG 10603  
6336 GCGCGGCGCATGCTGCTGCAACCGCATTAATAGCAAAATCGCGAGTACCGCAAAACA 6395  
10604 GCAGGGGCGCAACGCGCGCGCTGCGCAACGCGCATCAGCGCGATAGCCACCGCGCAACA 10663

QY 6396 CCACCGCAATAAGATCGTTGAGCTCAAACTTACCCTGTCGGTTCATGGTGCACAGAT 6455  
|||||  
Db 10664 CCACGCAAAAGATCGTTAGCTCAAAATACGCCCTTGGCGGGGTATGGTGTGACTCAT 10723  
|||||  
QY 6456 GCCAGCCCAATCCCAACCGTGCATGATGATTTATTCGACAGCCCGCTACGATTTCCA 6515  
|||||  
Db 10724 GCCAGCGGCATCCCAACCGCTGCATTAATGACGGTGGGTAAACGCGCGATGCCCTCCA 10783  
|||||

QY 6516 TCACCACACCGGTGTCACCAACAGATAA 6542  
|||||  
Db 10784 TCGCAATAACGCTCAAGATGACGATTA 10810  
|||||

RESULT 9  
US-10-474-536-45  
; Sequence 45, Application US/10474536  
; Publication No. US20040176570A1  
; GENERAL INFORMATION:  
; APPLICANT: Bachler, Adelbert  
; APPLICANT: Rohdich, Felix  
; APPLICANT: Adam, Petra  
; APPLICANT: Amslinger, Sabine  
; APPLICANT: Eisenreich, Wolfgang  
; APPLICANT: Hecht, Stefan  
; TITLE OF INVENTION: INTERMEDIATES AND ENZYMES OF THE NON-MEVALONATE ISOPRENOID  
; TITLE OF INVENTION: PATHWAY  
; FILE REFERENCE: 9286.24  
; CURRENT APPLICATION NUMBER: US/10/474,536  
; PRIORITY FILING DATE: 2003-10-10  
; PRIOR APPLICATION NUMBER: PCT/EP02/04005  
; PRIOR FILING DATE: 2002-04-10  
; PRIOR APPLICATION NUMBER: DE 102 01 458.2  
; PRIOR FILING DATE: 2002-01-16  
; PRIOR APPLICATION NUMBER: DE 101 55 084.7  
; PRIOR FILING DATE: 2001-11-09  
; PRIOR APPLICATION NUMBER: DE 101 30 236.3  
; PRIOR FILING DATE: 2001-06-22  
; PRIOR APPLICATION NUMBER: DE 101 18 166.3  
; PRIOR FILING DATE: 2001-04-11  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 45  
; LENGTH: 7494  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: pBScarot14  
US-10-474-536-45

Query Match 23.1%; Score 1618.8; DB 17; Length 7494;  
Best Local Similarity 66.0%; Pred. No. 0;  
Matches 2375; Conservative 0; Mismatches 1217; Indels 6; Gaps 2;  
QY 2490 CAGCAGCAGTGTGCTGGCGGAGCGACCTGATGCGACGCAATACAGATGATTTGG 2549  
|||||  
Db 3213 CACCGGTGAGCTGTCTTAAGTGGGAGCGGCTATGCAACCGCATATGATCTGATCTCG 3272  
|||||  
QY 2550 TCGGTGCTGAGCTGGCGAATGGCTTGATTCGCTGCGCTGCGTCAATTTGACGCCACAC 2609  
|||||  
Db 3273 TGGGGCTGAGCTCGCGAATGGCTTATCGCCCTGCGTCTTACGACGACGCAACCTGATA 3332  
|||||  
QY 2610 TGAATGCTGTGCTGGAGAGGATGCGCATCGCGACGCAATCATCTGCTGGTGTTC 2669  
|||||  
Db 3333 TGGGTATTTTGTCTATCGACGCGCGACCCAGCGCGGGGGAATCATACGTGGTCAATTC 3392  
|||||  
QY 2670 ATCAGACGATCTCAGCGCGCAACACTTCGCTGGCTGCAACCGCTGATTTACCGTGT 2729  
|||||  
Db 3393 ACCACATGATTTGATCTGAGACCAACATCGTTGGATAGCTCCGCTGGTGTTCATCAT 3452  
|||||  
QY 2730 GGTGAGGTTATCAGGTGCGTTTCTGCGCTGCGCGCAATCTGGACGGGGATTAATGTT 2789  
|||||  
Db 3453 GGCCCCGACTATCAGGTACGCTTTCCCAACGCGCTGTAAGCTGAACAGCGGCTACTTTT 3512  
|||||

QY 2790 CCATCGCATCAGCGCATTTTTCGCCGCCATCTTTTACGGCGCGATGGGTGACGATCTGTGGA 2849  
|||||  
Db 3513 GTATTACTTCTCAGCGTTTCGCTGAGGTTTACAGCGACAGTTCGCGCCGACATTTGTGGA 3572  
|||||  
QY 2850 CAAACACAGCGGTACACAGATTAACCCACGACGTGACGCTGCGGATGSCCGTGAAC 2909  
|||||  
Db 3573 TGGATACCGCGGTGCGAGAGGTAAATGCGGAATCTGTCGGTGAAGAGGTGAGGTTA 3632  
|||||  
QY 2910 TTGCTGCCCAAGTGGTGAATGATGTCGCGCGCTGACGCCGACGCCACATCTGACAGCTGG 2969  
|||||  
Db 3633 TCGGTGCCCGCGCGGTGATTCAGCGCGGGGTATTCGCGCAAAATTCAGCACTGAGCGTGG 3692  
|||||  
QY 2970 GTTATCAGGTGTTTCTTGGACAAGATGCGAGCTGCGCGACGCCGACGCCCTGACAGCAG 3029  
|||||  
Db 3693 GCTTCAGCGCTTTATTTGCCAGGAAATGCGCATTTGAGCCACCCGCAATGTTTATCGTCTC 3752  
|||||  
QY 3030 CGATCTGATGATGCCACCGTCGATCAGCAAGGGGTTATCGTTTCTCTACAGCTGTC 3089  
|||||  
Db 3753 CCAATTATCATGATGCCACCGTTCGATCAGCAAAATGTTTATCGCTTCGTGTACAGCTGTC 3812  
|||||  
QY 3090 CGCTCAGCGCGATCGGCTATTGATTTGAAGATACCCATTTAGCTTTAACAGCCCGCGCTGG 3149  
|||||  
Db 3813 CGCTCTCGCGACGATTTGTTAATTTGAAGACACGCACTATATTTGATAATTCGACATTAG 3872  
|||||  
QY 3150 CGGAAACACCGCTCTGTCAGCAATCGCGGCTATGCCAATCAGCAAGGCTGACGCTGA 3209  
|||||  
Db 3873 ATCTGAAATGCGCGCGCAAAATATTTGCGACTATGCGCGCAACAGGGTTGGCAGTTC 3932  
|||||  
QY 3210 GTACGCTGCTGCTGAAGACGACGCAATTTACCGATTTACCTGAGCGGCAACATCGATC 3269  
|||||  
Db 3933 AGACACTGCTGCGAGAGAACAGGGCGCTTACCCATTTACTCTGTCGGGCAATGCGCAGC 3992  
|||||  
QY 3270 GATTCTGGCAAACAGCAGCGCGCAAGCGTGCAGCGGCTGCGCGCGGCTGTTTCAATG 3329  
|||||  
Db 3993 CATTCTGG---CAGCAGCGCGCTGCTGAGTGATTTACGTCGCGCTCTGTTCCATC 4049  
|||||  
QY 3330 CCACCAACCGGTACTCTTCGCTGCGCTGCGGCTAGCGGCTGAGCGAGTTGTTAGCAGCGTGT 3389  
|||||  
Db 4050 TTACCAACCGGCTATTCTACTGCGCTGCGGTTGCGGCGACCGCTGAGTGACACTTG 4109  
|||||  
QY 3390 TGCCCAACCGATGCGCTCAGCTCAGCCCAACATATCGAACGCTTTTCCGCTCAGCAGTGGC 3449  
|||||  
Db 4110 ATGCTTTTACGTCGCGCTCAATTTACCATGCCATTTACGCAITTTGCCCGCGAGCTGGC 4169  
|||||  
QY 3450 GCGAAACAGCGATTTTTCGCTCTGTAAACCGCATGCTGTTTTTGGCGGTGAAGCCGACG 3509  
|||||  
Db 4170 AGCAGCAGGGCTTTTTCGCGATGCTGAATCGCATGCTGTTTTTTAGCCGCGACCCGCGATT 4229  
|||||  
QY 3510 AGCGCTGCGCGTGAATGCAACGTTTTTACCGCTCGATGCGGGTTAATTTAGCCGCTTTT 3569  
|||||  
Db 4230 CACGCTGCGGGTTATGACGCGTTTTTATGTTTACCTGAAGATTTAATTTGCCGCTTTT 4289  
|||||  
QY 3570 ACGCCGGCAACTGCGCTGCGCGATAAAACGCGGATTTCTGTGCGGCAAGCCCGCGTGC 3629  
|||||  
Db 4290 ATGCGGGAACACTCAGCTGACCGCATCGCTACGTTATTTCTGAGCGGCAAGCCGCTGTT 4349  
|||||  
QY 3630 CCAATCGGTGAAGCGCTGCGCGCTGTTGAAATCTGTGCAACCGGAGAGAAATAAGNA 3689  
|||||  
Db 4350 CGGTATTAGCAGCATTTGAAGCCATTATGACGACTCATCGTTTAAAGAGCGACTACATGAA 4409  
|||||  
QY 3690 AGCGCACTTATGTTGATTTGCGCAGGCTTTTGGCGCTGCGCTGCGGATTTTCGCTCAAGC 3749  
|||||  
Db 4410 ACCAACTACGGTAAATTTGTTGACGCTTTTTCGTTGCGGCTTGGCACTGCGCAATTCGTCTACAGC 4469  
|||||  
QY 3750 GCGCGGCATACCAACGACTTTACTCGACGACGCAACAAACGCGGCGGACGCGCTATGT 3809  
|||||  
Db 4470 TCGCGGGATCCCGCTTTACTGCTTGAACACGCTATAAACCCGCGGTCGCGGCTTATGT 4529  
|||||  
QY 3810 GTTTGAGACAGTGGCTTTACCTTCGATGCGCGGACCCAGGTGATCAGGATCCGAGCG 3869  
|||||  
Db 4530 CTACGAGGATCAGGGGTTTACCTTTGATGACAGGCCGCGGTTATCACCGATCCCGAGTGC 4589  
|||||  
QY 3870 CATCGAAGATTTGTTTCACTGCGTGGCAGGAAATTCGCTCAGCGATTTACGCTGAGCTGATGCC 3929  
|||||

Db 4590 CATGTGAAGAACTGTTTGCACTGCGCAGGAAACAGTTAAAGAGATATGTGAACTGCTGCC 4649  
Qy 3930 GGTAAACGCCCTCTATCGCTGTGCTGGGAAGATGGCAACAGCTTTGATTACGACAATAA 3989  
Db 4650 GGTGTACGCCGTTTACCGCTGTGTTGGGAGTCAGGGAAGTCTTTAAATACGATAACGA 4709  
Qy 3990 TCAGCGCTGCTGGAGCAGAGATCGCCAGTTTCAATCCGCAAGATGTAGAGGCTATCG 4049  
Db 4710 TCAACCCCGCTCGAAGCGCAGATTACGCAAGTTTAAATCCCGCGATGTGCGAAGTTATCG 4769  
Qy 4050 TCAATTTCTGCTCTATTCAGTGAAGTATTATGAGAGGGTTATCTGAACTCGGCACGGT 4109  
Db 4770 TCAGTTTCTGGACTATTACGCGCGGTGTTTAAAGAGGCTATCTAAAGCTCGTACTGT 4829  
Qy 4110 GCGTTTCTGAGTGCCTGACATGTGCGCGTTCGCGCGCAGTTGGGACGCTCGCAAGC 4169  
Db 4830 CCCTTTTATCGTTCAGAGACATGCTTCGCGCGCACCTCAACTCGCGAAACCTGCAGGC 4889  
Qy 4170 ATGGCGAGGCTCTACAGCATGTGTGGGAAATTTATTCAGGACGATCATCTGGTCAGGC 4229  
Db 4890 ATGGAGAAGCGTTTACAGTAAGTTGCGAGTTTACATCGAAGATGAACATCTGCGCCAGGC 4949  
Qy 4230 GTTTTCTTCCACTCATTTGCTGTGGCGGTGATGTTTCCGCGCGGCGCACCGCGC 4349  
Db 4950 GTTTTCTTCCACTCGTGTGTGGCGGCAATCCCTTCGCCACCTCATCTCATTTATAC 5009  
Qy 4290 CTTAATTCATGCGCTGGAGCGTAATGGGCGGTGTGTTTCCGCGCGGCGCACCGCGC 4349  
Db 5010 GTTGATACAGCGCTGGAGCGTGAGTGGGCGTCTGTTTCCGCGTGGCGCACCGCGC 5069  
Qy 4350 GCTGTGCAAGGATCGCGGACTGTTGAGGACTTGGGCGGAGCTGTTACTGAATGC 4409  
Db 5070 ATTAGTTCAGGGGATGAATAAGCTGTTTCAAGATCTGGGTGGCGAAGTGTGTTAAACGC 5129  
Qy 4410 CGAAGTGAGCCACTCGAAACACGCGCAATCGCAATTAGCGGTTCAGTTAGAGGCGG 4469  
Db 5130 CAGAGTCAGCCATATGGAACGACAGAAACAAGATTGAAGCGCTGCATTTAGAGACCG 5189  
Qy 4470 ACGACGCTTCGATGCGCGCGCTGTGGCCTCCAATGCGAGCTGGTGCATACCTTACGACAA 4529  
Db 5190 TCGCAGGTTCTTGACGCAAGCGTTCGCGTCAATGCGAGATGTTTACATACCTATCGCA 5249  
Qy 4530 ACTGCTTCGCGCAATTCGCTGGCAATGAACGTCGCGCAATCGCTGAAGCGTAAGCGCAT 4589  
Db 5250 CTTGTTAAGCAGCACCTTCCGCGGTAAAGCAGTCCAAACAACTGCAGACTAAGCGCAT 5309  
Qy 4590 GAGCAACTCGCTGTTGTACTCTATTTGCGCTGATCAGCGCATGAACAGCTCGCGCA 4649  
Db 5310 GAGTAACTCTGTGTTGTCTCTATTTGTTTGAATCACCATCATGATCAGCTCGCGCA 5369  
Qy 4650 CCACACCGCTGTGTTTGGCGCGGTTATCGTGAAGTTGATCGATGAGATTTTCAACAGCAG 4709  
Db 5370 TCACACGGTTGTTTCGCGCGCGTTTACCGGAGCTGATTGACGAAATTTTAAATCATGA 5429  
Qy 4710 CCGAGTCGAGAGGATTTTCACTTACCTGACGCGCGCTGAGCAGGAGATCCGTCGCT 4769  
Db 5430 TGGCGCTCGAGAGGACTTCTCACTTTATCTGACGCGCGCTGTGTCAAGGATTCGTCAT 5489  
Qy 4770 GGCACCGCGCTGCGGCGAGCTTTTATGTTGTAGCGCGGTGCGCATCTCGGCACCGC 4829  
Db 5490 GCGCGCTGAAGGTTGCGCGAGTTACTATGTTGTGGCGCGGTGCCCATTTAGGCAACCGC 5549  
Qy 4830 TGACATCGACTGCGCAACAGGAAGGACCGCGCTTTCGCGGATCGAAATTTTTCGTTATCTGA 4889  
Db 5550 GAACCTCGACTGGAACGGTTGAGGGGCAAACTACCGGACCGTATTTTTCGCTACCTTGA 5609  
Qy 4890 GAGCACTACATGCGCGGATTTAGTCAGCAATTTAGTGAACACAGAAATTTTACGCGCTT 4949  
Db 5610 GCAGCAATTACATGCTGGCTTACGAGTCAAGTGTGTCAGCACCGGATGTTTACGCGCTT 5669  
Qy 4950 TGATTTTCCGCGACACGCTGCATGCCCATCAGGCTCGGCGTTTTCGCTGAGACCGGATTTT 5009

Db 5670 TGATTTTCCGACACAGCTTAATGCGCTATCATGGCTCAGCGCTTTTCTGTGAGAGCCCGTTCT 5729  
Qy 5010 GAGCGAAAGCGCTGTGTTCCGCCCGCATAAACCGCGATGCGGATATCAGCAATCTCTATCT 5069  
Db 5730 TACCCAGAGCGCTGTGTTTCGCGCGCATTAACCGGATAAAACCATTAATCTCTACCT 5789  
Qy 5070 GGTGGTGGCGGTACGATCAGCGCGGCGGTGCGCGCGGTGATCGGTTGCGGCAAGGC 5129  
Db 5790 GGTGGCGGAGCAGCATCCGCGGCGAGGCATTCCTGCGTCACTCGGCTCGGCAAAAGC 5849  
Qy 5130 CACCGCAGGCTGATCTCGAGGATCGCGCGCAATGAATCGACAGCTTTTACTTTAGCAA 5189  
Db 5850 GACAGCAGGTTTGTGCTGAGGATC---TGATTTGAATTAATCCGTCGTTACTCAATCAT 5906  
Qy 5190 GTAACGCAAAACCATCGCTGGGCTGGAAGAGTTTCGCCACCGCCCGCAAGCTGTTTGAT 5249  
Db 5907 GCGGTGCAAAACGATGCGAGTTGGCTCGAAAGTTTTTGCACAGCCTCAAAAGTTATTGAT 5966  
Qy 5250 GCACGACGCGCGGACGACGCTGATGCTGTATGCGTGTGTGCTCACTCGGATGATGTG 5309  
Db 5967 GCAAAACCCCGCGCAGCGTACTGATGCTTACGCTGTGTGCGGCAATTTGTGACGATGT 6026  
Qy 5310 ATTGATGGCAAAACGCTGGCGGAAGCGGCGACGAGCATGCGTTCGAAGACGCGCAGGCA 5369  
Db 6027 ATTGACGATCAGACGCTGGGCTTTTTCAGGCCCGGAGCCTGCTTACAAACGCCCGCAACA 6086  
Qy 5370 CGTATGACGATCTGCAAAATGAAACCCCGCGGCTTACAGCGCGCGCGCACATGAGTGA 5429  
Db 6087 CGTCTGATCAACTTGAGATGAAACCGCGCCAGGCTATGCGAGGATCGCAGATGACGAA 6146  
Qy 5430 CCGCGGTTTAGGCGGTTTTCAGGAAGTGGGATCATTCACAGCTGCGGCAACACTGGCG 5489  
Db 6147 CCGCGGTTTTCGCGCTTTTTCAGGAAGTGGCTATGGCTCATGATATCGCCCGCTTACGCG 6206  
Qy 5490 TTTGATCATCTCGAAAGCTTTCGCTATGATGACGCAACGAAATTCACGAGCTTCGAT 5549  
Db 6207 TTTGATCATCTGGAAGGCTTCGCCATGATGATGCGGAAGCGCAATACAGCCAACTGGAT 6266  
Qy 5550 GACAGCTGCGTTACTGCTATCACGTCGCGGCGTGGTTCGTTTGTATGATGGCGCGCTGA 5609  
Db 6267 GATACGTCGCTATTTGCTATCACGTTGCGAGCGTTGTCGCTTGTATGATGGCGCAATC 6326  
Qy 5610 ATGGGCGTGGCGACGAAGCGGTGCTCGATCAGCGCTGCGATTTAGGACTGGGTTCCAG 5669  
Db 6327 ATGGGCGTGGGGAATAACGCCACGCTGGACCGCGCTGTGACCTTGGGCTGGCAATTCAG 6386  
Qy 5670 CTCACATACTTCGCGCGACATTTAGAAAGATGCGGAAATGGTTCGCTGCTATCTGCGC 5729  
Db 6387 TTGACCAATATTGCTCGCATATTGTTGACGATGCGCATGCGGCGCGCTGTTATCTGCGC 6446  
Qy 5730 CAATCTGCTCGATCAGCGGGGATTACGCGCGGATACGCTGACCTGACCGCAACATCGT 5789  
Db 6447 GCAAGCTGCTGAGCATGAAGGCTGAACAAAGAGAAATTAATGCGGCACTGGAACCGT 6506  
Qy 5790 GACGCGCTCGCTCAGTGGCAGCGGTATTAGTGGGAGGCGGAAACCTTATTATCACTCG 5849  
Db 6507 CAGGCGCTGAGCGGTATCGCCGCTGTTTGGTGCAGGAGCAGAACTTACTATTGTTCT 6566  
Qy 5850 GCGCATCCGCTTACCGGTTTACCGCTGCGCTGGCGTGGCGCATCGCTACGCGCTCGC 5909  
Db 6567 GCCACAGCGGCTGGCAGGTTGCCCTGCGTTCCGCTGGGCAATCGCTACGCGGAAG 6626  
Qy 5910 GCGGTTTATCGCGAAATTTGGCGTCAAAGTTTCAGCAGCGCGGTGTGACCGCTGGGATTC 5969  
Db 6627 CAGGTTTACCGGAAATAGTGTCAAAGTTTGAACAGGCGGTCAGCAAGCCTGGGATCAG 6686  
Qy 5970 CCGCAGCGCACAGTAAAGGTGAAAACTGGGCGCTGCTGGTGAAGGGGCGAGGTTTGGCG 6029  
Db 6687 CCGCAGTCAACACCAACCGCCCGAAAAATTAACGCTGCTGCTGCGCGCTCTGCTCAGGCC 6746  
Qy 6030 ATCACTTCGCGTGTCTCGTCTGAAACCGCGTTCGCGTGGTCTGTGGCAGGCTCTC 6087  
Db 6747 CTTACTTCCCGATCGGCGCTCATCTCCCGCGCTGCGCATCTCTGGCAGCGCCCGC 6804









FILE REFERENCE: CL2360 US NA  
CURRENT APPLICATION NUMBER: US/10/808,979  
CURRENT FILING DATE: 2004-03-25  
PRIOR APPLICATION NUMBER: US 60/471,904  
PRIOR FILING DATE: 2003-05-20  
PRIOR APPLICATION NUMBER: US 60/527,083  
PRIOR FILING DATE: 2003-12-03  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 7  
LENGTH: 1482  
TYPE: DNA  
ORGANISM: Enterobacteriaceae strain DC260  
US-10-808-979-7

Query Match 21.2%; Score 1482; DB 18; Length 1482;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

3685 ATGAACGCACTTATGTGATTGGCGCAGGCTTTGGCGGCTGGCGCTGGCGATTGGCGCTG 3744  
1 ATGAACGCACTTATGTGATTGGCGCAGGCTTTGGCGGCTGGCGCTGGCGATTGGCGCTG 60  
3745 CAAGCGCGGGCATACCAACACCTTATCGAGCAGCGCAAAACCGGCGGAGCGGCC 3804  
61 CAAGCGCGGGCATACCAACACCTTATCGAGCAGCGCAAAACCGGCGGAGCGGCC 120  
3805 TATGTGTTTGGAGCAGTGGCTTTACCTTCGATGCGGACCCACCGGTGATCCAGATCCC 3864  
121 TATGTGTTTGGAGCAGTGGCTTTACCTTCGATGCGGACCCACCGGTGATCCAGATCCC 180  
3865 AGCGCCATCGAAGTGTTCACGCTGGCAGGAAATCCCTCAGCGATTACGTCGAGCTG 3924  
181 AGCGCCATCGAAGTGTTCACGCTGGCAGGAAATCCCTCAGCGATTACGTCGAGCTG 240  
3925 ATGCCGGTAACGCCCTTCTATGCTGTGTGGGAAAGATGGCAACAGCTTGATTACGAC 3984  
241 ATGCCGGTAACGCCCTTCTATGCTGTGTGGGAAAGATGGCAACAGCTTGATTACGAC 300  
3985 AATAATCAGCGCTGTCGAGCAGCAGATCGCCAGCTTCAATCCGCAAGATGTAGAAGGC 4044  
301 AATAATCAGCGCTGTCGAGCAGCAGATCGCCAGCTTCAATCCGCAAGATGTAGAAGGC 360  
4045 TATCGTCAATTTCTTCCCTATTACGCTGAAGTATTTAGAGAGGTTTATCGAACTCGGC 4104  
361 TATCGTCAATTTCTTCCCTATTACGCTGAAGTATTTAGAGAGGTTTATCGAACTCGGC 420  
4105 ACGGTGCGGTTTCTGACAGTGTGACATGCTCGCGTGGCGCGCAGTTGGAGCTGCTG 4164  
421 ACGGTGCGGTTTCTGACAGTGTGACATGCTCGCGTGGCGCGCAGTTGGAGCTGCTG 480  
4165 CAAGCATGGCGCAGCTGTACAGCATGGTGGGAAATTTATTTCAGGACCATCATCTGCGT 4224  
481 CAAGCATGGCGCAGCTGTACAGCATGGTGGGAAATTTATTTCAGGACCATCATCTGCGT 540  
4225 CAGCGGTTTCTTCCATCATTTGCTGGTGGCGGTAATCTTTTCAACAGTCATCGATC 4284  
541 CAGCGGTTTCTTCCATCATTTGCTGGTGGCGGTAATCTTTTCAACAGTCATCGATC 600  
4285 TATACCTTAATTCATGCGCTGGAGCGTGAATGGGCGTGTGGTTTCCGCGCGCGGCACC 4344  
601 TATACCTTAATTCATGCGCTGGAGCGTGAATGGGCGTGTGGTTTCCGCGCGCGGCACC 660  
4345 GGGCGCTGTGTCAGGCGCATGGCGCATGTTTCAGGACTTGGGCGCGAGCTGTTACTG 4404  
661 GGGCGCTGTGTCAGGCGCATGGCGCATGTTTCAGGACTTGGGCGCGAGCTGTTACTG 720  
4405 AATGCCGAAGTAGCAGCTGGAAACAGCGGCAATCGCAATTCAGCGGCTTCAGTAGAG 4464  
721 AATGCCGAAGTAGCAGCTGGAAACAGCGGCAATCGCAATTCAGCGGCTTCAGTAGAG 780  
4465 GCGGCGAGCGCTTCGATGCGCGCTTCCAAATCGCCAGCTGGTGCATACCTAC 4524

Db 781 GCGGACGACGCTTTCGATGCGCGCTGTGGCTCCCAATGCGAGCTGGTGATACCTAC 840  
Qy 4525 GACAACTGCTTTCGCCACCATCCGCTGCAATGAAACGCTGACATCGCTCAAGCGTAAG 4584  
Db 841 GACAACTGCTTTCGCCACCATCCGCTGCAATGAAACGCTGACATCGCTCAAGCGTAAG 900  
Qy 4585 CGCATGAGCAACTCGCTGCTTTGTACTCTATTTTGGCTTGAATCAGCCGCATGAACAGCTC 4644  
Db 901 CGCATGAGCAACTCGCTGCTTTGTACTCTATTTTGGCTTGAATCAGCCGCATGAACAGCTC 960  
Qy 4645 GCGCACACACGCTGCTTTTGGCGCGGTTATCGTGAGTTGATCGATGAGATTTTCAAC 4704  
Db 961 GCGCACACACGCTGCTTTTGGCGCGGTTATCGTGAGTTGATCGATGAGATTTTCAAC 1020  
Qy 4705 AGCAGCAGCTGGCAGACGATTTTTCATTTTACCTTGACACGCGCCCTGCGAGCGATCCG 4764  
Db 1021 AGCAGCAGCTGGCAGACGATTTTTCATTTTACCTTGACACGCGCCCTGCGAGCGATCCG 1080  
Qy 4765 TCGTGGCACCGCGCGCTGCGGAGCTTTTATGTGTTAGCGCGGTGCCGATCTCGGC 4824  
Db 1081 TCGTGGCACCGCGCGCTGCGGAGCTTTTATGTGTTAGCGCGGTGCCGATCTCGGC 1140  
Qy 4825 ACCGCTGACATCGACTGCGACAGGAGCGCGCTTGGCGATCGAATTTTTCCTTAT 4884  
Db 1141 ACCGCTGACATCGACTGCGACAGGAGCGCGCTTGGCGATCGAATTTTTCCTTAT 1200  
Qy 4885 CTGGAGCAGCACTACATCCCGGATTACCTCAGCAATTTAGTGACACAGAAATGTTTACG 4944  
Db 1201 CTGGAGCAGCACTACATCCCGGATTACCTCAGCAATTTAGTGACACAGAAATGTTTACG 1260  
Qy 4945 CGGTTTGAATTTTGGCACACGCTGATCCCATCAAGCTCGGCTGGCGTTTTCGCTGAGCGG 5004  
Db 1261 CGGTTTGAATTTTGGCACACGCTGATCCCATCAAGCTCGGCTGGCGTTTTCGCTGAGCGG 1320  
Qy 5005 ATTTTGAAGCAAGCGCTGTTCCGCGCGATACCCGATGCGATATCAGCAATCTC 5064  
Db 1321 ATTTTGAAGCAAGCGCTGTTCCGCGCGATACCCGATGCGATATCAGCAATCTC 1380  
Qy 5065 TATCTGTGGTGGTGGCTGACGATCCAGCGCGGCGGTGCCGCGGTGATCGGTTTCGCGC 5124  
Db 1381 TATCTGTGGTGGTGGCTGACGATCCAGCGCGGCGGTGCCGCGGTGATCGGTTTCGCGC 1440  
Qy 5125 AAGGCGCGCGCGCGCTGATGCTGAGGATCGCGCGGATGA 5166  
Db 1441 AAGGCGCGCGCGCGCTGATGCTGAGGATCGCGCGGATGA 1482

RESULT 12  
US-10-808-979-3  
Sequence 3, Application US/10808979  
Publication No. US20040268439A1  
GENERAL INFORMATION:  
APPLICANT: E.I. duPont de Nemours and Company, Inc.  
APPLICANT: Cheng, Qiong  
APPLICANT: Tao, Luan  
TITLE OF INVENTION: GENES ENCODING CAROTENOID COMPOUNDS  
FILE REFERENCE: CL2360 US NA  
CURRENT APPLICATION NUMBER: US/10/808,979  
CURRENT FILING DATE: 2004-03-25  
PRIOR APPLICATION NUMBER: US 60/471,904  
PRIOR FILING DATE: 2003-05-20  
PRIOR APPLICATION NUMBER: US 60/527,083  
PRIOR FILING DATE: 2003-12-03  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 3  
LENGTH: 1278  
TYPE: DNA  
ORGANISM: Enterobacteriaceae strain DC260  
US-10-808-979-3

Query Match 18.3%; Score 1278; DB 18; Length 1278;  
Best Local Similarity 100.0%; Pred. No. 0;

Matches 1278; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1245 GTGGCCACTTTGCGGTGATCGCGCGCGCTCTACAGCCACTTTTCA CGGTTCGAGGCG 1304
DB	1 GTGGCCACTTTGCGGTGATCGCGCGCGCTCTACAGCCACTTTTCA CGGTTCGAGGCG 60
QY	1305 TTAGCAAAACGCTGCTGGGCGCGGCGATCGCATCACATTATCAGCAGACCGGATGCC 1364
DB	61 TTAGCAAAACGCTGCTGGGCGCGGCGCATCGCATCACATTATCAGCAGACCGGATGCC 120
QY	1365 CGCACTTTGTAGCGAACGAACCATCGATTTTGTGTCGTGGCGCCAAACAGACGCATCCT 1424
DB	121 CGCACTTTGTAGCGAGCAACGCATCGATTTTGTGTCGTGGCGCCAAACAGACGCATCCT 180
QY	1425 GCCGGTTTCGTGGCGCCCGGTGTGTCATCGGCTGGCCCTCGCGGGCGGCTGTGCCTGTTT 1484
DB	181 GCCGGTTTCGTGGCGCCCGGTGTGTCATCGGCTGGCCCTCGCGGGCGGCTGTGCCTGTTT 240
QY	1485 CGGCTGATCGACGATCTCGGCTCTGCAACGATATGCTGTGCCGGAACCTGCTCGGTA 1544
DB	241 CGGCTGATCGACGATCTCGGCTCTGCAACGATATGCTGTGCCGGAACCTGCTCGGTA 300
QY	1545 CTCGAAAGCATTTGAACATCGATGCGGTGATCGCGAGCAAAATGAAGCGGCGGCGGATTCG 1604
DB	301 CTCGAAAGCATTTGAACATCGATGCGGTGATCGCGAGCAAAATGAAGCGGCGGCGGATTCG 360
QY	1605 GTCGCTGAAGCGCTGCATCTGCCGTTTGTGTTTCGGTGGCCTCGCCCTTGC CGGTCAATCGT 1664
DB	361 GTCGCTGAAGCGCTGCATCTGCCGTTTGTGTTTCGGTGGCCTCGCCCTTGC CGGTCAATCGT 420
QY	1665 GAAGCGGGGANTCCGCTTGGCGGTGATGCCCTTCGTTTGTGCA CAGGATGACAAGCGCTG 1724
DB	421 GAAGCGGGGANTCCGCTTGGCGGTGATGCCCTTCGTTTGTGCA CAGGATGACAAGCGCTG 480
QY	1725 AAACGTTTTCAGGCCACGACGGATCTATGATCGCATCATGCTGCTGTCACGGGACGTG 1784
DB	481 AAACGTTTTCAGGCCACGACGGATCTATGATCGCATCATGCTGCTGTCACGGGACGTG 540
QY	1785 ATCCTCAAAACACGCGGGCGTTTAATTTGACGGAGCGGCGGATTTACATCAGTGCCTG 1844
DB	541 ATCCTCAAAACACGCGGGCGTTTAATTTGACGGAGCGGCGGATTTACATCAGTGCCTG 600
QY	1845 TCSCCGCTGSCACAAATCAGCCAGATGGTGC GGGCCTTTGATTTTCCATCGCAACTG 1904
DB	601 TCSCCGCTGSCACAAATCAGCCAGATGGTGC GGGCCTTTGATTTTCCATCGCAACTG 660
QY	1905 CCCGCCCTGCTATCACCGCGTGGGGCCACTCGCGCCCCGGTTTCTCCTCGCGCGCTCCAT 1964
DB	661 CCCGCCCTGCTATCACCGCGTGGGGCCACTCGCGCCCCGGTTTCTCCTCGCGCGCTCCAT 720
QY	1965 GCGCCTTGGCAGCGCTCGCTCAGCCGGTGTTATGCCCTCGCTGGGTGACGCTGCAAGGC 2024
DB	721 GCGCCTTGGCAGCGCTCGCTCAGCCGGTGTTATGCCCTCGCTGGGTGACGCTGCAAGGC 780
QY	2025 CATCGCTTCGGGCTGTTTCTGTCATCTGGCGACGCGTGC CGCCAGCTGCGGCTATCGCTG 2084
DB	781 CATCGCTTCGGGCTGTTTCTGTCATCTGGCGACGCGTGC CGCCAGCTGCGGCTATCGCTG 840
QY	2085 GTGATCGCCCATTTGTGGGGGATTTAAACGCCGGAACAGACGATCAGCTGGAGCTGCGTGCC 2144
DB	841 GTGATCGCCCATTTGTGGGGGATTTAAACGCCGGAACAGACGATCAGCTGGAGCTGCGTGCC 900
QY	2145 GCGCGTGGGTGACGGAATTTCTGTCATCTGATCAGCGCGACCCCTPACAGCACGCGCAGCTGTTT 2204
DB	901 GCGCGTGGGTGACGGAATTTCTGTCATCTGATCAGCGCGACCCCTPACAGCACGCGCAGCTGTTT 960
QY	2205 ATCACTCATGCGGGTTTAAACAGCGCGCTGGAAGCACTGGAATTCGGTACGCCGATCGTG 2264
DB	961 ATCACTCATGCGGGTTTAAACAGCGCGCTGGAAGCACTGGAATTCGGTACGCCGATCGTG 1020
QY	2265 GCGCTGCCGATTTGCTTTTGATTCAGCCCGGCGTGGCGGGCGCATTTGATGGCATGACGTT 2324
DB	1021 GCGCTGCCGATTTGCTTTTGATTCAGCCCGGCGTGGCGGGCGCATTTGATGGCATGACGTT 1080

Qy	2325	GGTGGCGCGCATCAGCGTTTAGCGGTTCATCACTGGAGCAGCATCTGCAACAGCTG	2384
Db	1081	GGTGGCGCGCATCAGCGTTTAGCGGTTCATCACTGGAGCAGCATCTGCAACAGCTG	1140
Qy	2385	CTGACCGACCATCGTTACGGCGCTACGGATGTACGGATTCAGGCGCAGCTGCAAGTCTG	2444
Db	1141	CTGACCGACCATCGTTACGGCGCTACGGATGTACGGATTCAGGCGCAGCTGCAAGTCTG	1200
Qy	2445	GGCGGTTCGACGCGTGGCGCGGCATCGTCCAGCAGCGCGCTGTGCCAGCAGCAAGTCTG	2504
Db	1201	GGCGGTTCGACGCGTGGCGCGGCATCGTCCAGCAGCGCGCTGTGCCAGCAGCAAGTCTG	1260
Qy	2505	CTGCGGAGCGCGACCTGA	2522
Db	1261	CTGCGGAGCGCGACCTGA	1278

RESULT 13  
 US-10-808-979-5  
 ; Sequence 5, Application US/10808979  
 ; Publication No. US20040268439A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: E.I. duPont de Nemours and Company, Inc.  
 ; APPLICANT: Cheng, Qiong  
 ; APPLICANT: Tao, Luan  
 ; TITLE OF INVENTION: GENES ENCODING CAROTENOID COMPOUNDS  
 ; FILE REFERENCE: CL2360 US NA  
 ; CURRENT APPLICATION NUMBER: US/10/808,979  
 ; CURRENT FILING DATE: 2004-03-25  
 ; PRIOR APPLICATION NUMBER: US 60/471,904  
 ; PRIOR FILING DATE: 2003-05-20  
 ; PRIOR APPLICATION NUMBER: US 60/527,083  
 ; PRIOR FILING DATE: 2003-12-03  
 ; NUMBER OF SEQ ID NOS: 20  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 5  
 ; LENGTH: 1167  
 ; TYPE: DNA  
 ; ORGANISM: Enterobacteriaceae strain DC260  
 ; US-10-808-979-5

Query Match	16.7%;	Score 1167;	DB 18;	Length 1167;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1167;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	2522	ATGCGCAGCAATACGATGTGATTTTGGTCGGTCTGGACTGGCGAATGGCTTGATTTGCG	2581	
DB	1	ATGCGCAGCAATACGATGTGATTTTGGTCGGTCTGGACTGGCGAATGGCTTGATTTGCG	60	
QY	2582	CTGCGTCTGGCTCAATTTGCAGCCACAACTGAAATGCCCTGTTCGTGGAGAGCGATGCGCAT	2641	
DB	61	CTGCGTCTGGCTCAATTTGCAGCCACAACTGAAATGCCCTGTTCGTGGAGAGCGATGCGCAT	120	
QY	2642	CCGCGAGCGCAATCATACCTGGTGTTTCATCACAGCGATCTCAGCGCGCAACCACTTCGCG	2701	
DB	121	CCGCGAGCGCAATCATACCTGGTGTTTCATCACAGCGATCTCAGCGCGCAACCACTTCGCG	180	
QY	2702	TGGCTGCAACCGCTGATTACCGTSCGTTGGTCAGGTTATCAGGTGCGTTTCTCGCGCTG	2761	
DB	181	TGGCTGCAACCGCTGATTACCGTSCGTTGGTCAGGTTATCAGGTGCGTTTCTCGCGCTG	240	
QY	2762	CGCGGCAATCTGGACGGGGATTAATGTTCCATCGCATCAGGCGAATTTTGGCCGCGACTTT	2821	
DB	241	CGCGGCAATCTGGACGGGGATTAATGTTCCATCGCATCAGGCGAATTTTGGCCGCGACTTT	300	
QY	2822	TACCGCGCGATGGGTGACGATCTCTGGACACAAACACAGCCGTACAAAGGTAAAAACCCACG	2881	
DB	301	TACCGCGCGATGGGTGACGATCTCTGGACACAAACACAGCCGTACAAAGGTAAAAACCCACG	360	
QY	2882	CAGGTGACGCTGGCGGATGGCCGTGAACCTTGCTCGCAAGTGGTGATTTGATGTCGCGGC	2941	
DB	361	CAGGTGACGCTGGCGGATGGCCGTGAACCTTGCTCGCAAGTGGTGATTTGATGTCGCGGC	420	

Qy	2942	CTGCAGCCGACGCCACACTCTGCAGCTGGGTTATCAGGTGTTTCTTGACACAGAGTGCAG	3001
Db	421	CTGCAGCCGACGCCACATCTGCAGCTGGGTTATCAGGTGTTTCTTGACACAGAGTGCAG	480
Qy	3002	CTGGCGCAGCCGACGGGCTTCAGCAGCCGATCTTGATGGATGCCACCGTCGATCAGCAA	3061
Db	481	CTGGCGCAGCCGACGGGCTTCAGCAGCCGATCTTGATGGATGCCACCGTCGATCAGCAA	540
Qy	3062	CGCGGTTATCGTTTTGTCTACACGCTGCGCTCAGCGCGGATCGGCTATTGATTGAAGAT	3121
Db	541	CGCGGTTATCGTTTTGTCTACACGCTGCGCTCAGCGCGGATCGGCTATTGATTGAAGAT	600
Qy	3122	ACCCATTACGTTTAACCAAGCCGCGCTGCGGAGAACCCGCTCGTCAGCACATCGCCGAC	3181
Db	601	ACCCATTACGTTTAACCAAGCCGCGCTGCGGAGAACCCGCTCGTCAGCACATCGCCGAC	660
Qy	3182	TATGCCAATCAGCAAGGCTGGAAGCTGAGTACGCTGCTGCGTGAAGACACGGCATATTA	3241
Db	661	TATGCCAATCAGCAAGGCTGGAAGCTGAGTACGCTGCTGCGTGAAGACACGGCATATTA	720
Qy	3242	CCGATTACCTGAGCGGCACATCGATCGATTCTGGCAAACAGCAGCGCGCCAGCGTGC	3301
Db	721	CCGATTACCTGAGCGGCACATCGATCGATTCTGGCAAACAGCAGCGCGCCAGCGTGC	780
Qy	3302	AGCGECCTGGCGCGCGGCTGTTTCATGCCCCACACCGGTTACTTCTGCGCTCGCGCGTG	3361
Db	781	AGCGECCTGGCGCGCGGCTGTTTCATGCCCCACACCGGTTACTTCTGCGCTCGCGCGTG	840
Qy	3362	CGCTTAGCGAGTTGTTAGCAGCGCTTTGCCCCACCGATGCCCTCAGCTCAGCCCAACAT	3421
Db	841	CGCTTAGCGAGTTGTTAGCAGCGCTTTGCCCCACCGATGCCCTCAGCTCAGCCCAACAT	900
Qy	3422	ATCGAAGCTTTGCGCGCTCAGCAGTGGCGGAACAGCGATTTTTCGCTCTGCTAAACCGC	3481
Db	901	ATCGAAGCTTTGCGCGCTCAGCAGTGGCGGAACAGCGATTTTTCGCTCTGCTAAACCGC	960
Qy	3482	ATGCTGTTTTGGCCGGTAAAGCCGACAGCGCTGGCGCGTGATGCAACGTTTTTACCGG	3541
Db	961	ATGCTGTTTTGGCCGGTAAAGCCGACAGCGCTGGCGCGTGATGCAACGTTTTTACCGG	1020
Qy	3542	CTCGATGCCGGGTTAATTAGCCGTTTTTACCGGGCAACTGCGCTGCGCGATATAAACG	3601
Db	1021	CTCGATGCCGGGTTAATTAGCCGTTTTTACCGGGCAACTGCGCTGCGCGATATAAACG	1080
Qy	3602	CGGATTCGTGCGCAAGCCCGGCTGCCCATCGGTGAAGCGCTGCGCGCTGTTTGAAT	3661
Db	1081	CGGATTCGTGCGCAAGCCCGGCTGCCCATCGGTGAAGCGCTGCGCGCTGTTTGAAT	1140
Qy	3662	TCGTGTGAACCGAGGGAAGAAAAATGA	3688
Db	1141	TCGTGTGAACCGAGGGAAGAAAAATGA	1167

```

RESULT 14
US-10-804-677-7
; Sequence 7, Application US/108046577
; Publication No. US20040224383A1
; GENERAL INFORMATION:
; APPLICANT: E.I. duPont de Nemours and Company, Inc.
; APPLICANT: Cheng, Qiong
; APPLICANT: Tao, Luan
; APPLICANT: Sedkova, Natalia
; TITLE OF INVENTION: GENES ENCODING CAROTENOID COMPOUNDS
; FILE REFERENCE: CL2346 US NA
; CURRENT APPLICATION NUMBER: US/10/804,677
; CURRENT FILING DATE: 2004-03-19
; PRIOR APPLICATION NUMBER: US 60/468,596
; PRIOR FILING DATE: 2003-05-07
; PRIOR APPLICATION NUMBER: US 60/527,083
; PRIOR FILING DATE: 2003-12-03
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2

```

```

; SEQ ID NO 7
; LENGTH: 1482
; TYPE: DNA
; ORGANISM: Pectobacterium cyripedii DC416
US-10-804-677-7

Query Match      15.5%; Score 1085.2; DB 18; Length 1482;
Best Local Similarity 83.3%; Pred. No. 0;
Matches 1234; Conservative 0; Mismatches 248; Indels 0; Gaps 0;

QY 3685 ATGAACCGCACTTATGTGATTTGGCGCAGGCTTTGGCGGCGCTGGCGCTGGCGATTCGCCTG 3744
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 ATGAACCGCACTTATGTGATTTGGCGCAGGCTTCGTGGCTGGCGATTCGCCTG 60

QY 3745 CAAGCGCGGGGCATACCAACCACTTACTCGACAGCGGACAAACCGGCGGACGCGCC 3804
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 CAAGCGCGCGCGTGCCTGTCAGCTGCTGAAACAGCGGATTAAGCTGCGCGGCGCGCC 120

QY 3805 TATGTGTTTCAAGACAGTGTCTTACCTTCGATGCCGACCCACGCTGATCACCGATCCC 3866
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 121 TATGTGTATCAGATCAGGGTTTTACCTTTGATGCCGGTCCGACGGTGATTCAGATCCC 180

QY 3865 AGGCCCATCGAAGATTGTTTCACTGTGCGAGGAAAAATCGCTCAGCGATTACGTCGAGCTG 3924
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 181 AGGCGTATCAGGCGCTGTTTACGCTGGCAGGCAACAATCACTGATGATTCGACCTG 240

QY 3925 ATCCGGTACGCGCTTCTATCGCTGTGCTGGAGATGGCNAACAGCTTGNATTACGAC 3984
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 241 ATCCGGTGACGCCATTTTATCGCCTGTGCTGGGAAGCGGACGAGCTGGACTACGAC 300

QY 3985 AATAATCAGCGCGTGTGGAGCAGCAGATCGCCACCTTCAATCCGCAAGATGTAGAAGC 4044
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 301 AACCAATCAGCGCAGCTGGAGCAGCAGATGCACTTTTAATCCCAGGATGTCGCGGT 360

QY 4045 TATCGTCAATTTCTTGCTATTTCACGTGAAGTATTTAGAGAGGTTATCTGAAAATCTCGC 4104
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 361 TACCGCGAGTTTCTGCGCTATTACAGAGATGTGTTTCTGAGGGGCTATCTGAAAATCTGGC 420

QY 4105 ACGGTGCGGTTTCTGCAAGTGCATGCTGCGCGTCCGCGCAGTTGGAGCTGTG 4164
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 421 ACGGTACCTTTCTGCAATTTCCGCGACATGCTGCGTGCCGGCCACAGCTGGGTGGCTG 480

QY 4165 CAAGCATGGCGCAGCGTCTACAGCATGGTGGGAAATTTATTCAGGACGATCATCTGGCT 4224
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 481 CAGCGCTGGCGCAGTGTCTACAGCATGGTGGGAAATTTATTCATGACGATCATCTGCGC 540

QY 4225 CAGCGGTTTTCTTCCACTCATGCTGGTGGCGGTAACTCTTTGCAACGTCATCGATC 4284
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 541 CAGGCTTTTCTCTTCACTGCTGTGCTGGCGGTAATCTTTTGCACAGCTTCTGATC 600

QY 4285 TATACCTTAATTCATGCGCTGGAGCGTGAATGGGCGTGTGTTTCCGCGCGCGGCACC 4344
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 601 TATACCTTAATTCACGCACTGGAGCGGAAATGGGCGGTGTGTTTCCGCGCGCGGTACC 660

QY 4345 GCGCGCGTGGTCAGGGGATGGCGACGTGTCAGAGACTTGGGCGGCGAGCTGTGTACTG 4404
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 661 GGTGCGCTGGTTGATGGCATGGCGCGGCTGTTTTCGCGATTTTGGGCGGTGAACTGTGCTC 720

QY 4405 AATGCCAAGTGAAGCAGCTGGAACACGCGGCAATTCGATTTAGCGGGCTTCAGTTTAGAG 4464
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 721 AACGCCAAGTCAACGACTGGACCTGGAGACCGAGGTTAACCGCATCAGCGGTGTCAGCTGAAG 780

QY 4465 GCGGGAAGACGCTTCGATGCGCGCGCTGTGGCTCCAATGCGGACGTGTGATACCTTAC 4524
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 781 GATGGGCGCGGTTTTTCCCGCGCGCGTGGGTCAAAATGCTGAAGTGTGTCATACCTTAC 840

QY 4525 GACAAATCGCTTCGCCACCATCGCTGGCAATGAACGCTGCCACATCGCTGAAGCGTGAAG 4584
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 841 GATCGCTGTTAAGCAGCATCTCGCGCGCTTAAACGCGCGGCAACGCTGAAGCGCAAG 900

QY 4585 CGCATGACCAACTCGCTGTTTGTACTCTATTTTGGCGCTGAATCAGCCGCGATGAACAGCTC 4644
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 901 CGATGAGCAACTCGCTGTTTGTACTCTATTTTGGTCTTAATCATGCCCCACCGCAGCTG 960

```









QY 5726 GCGCAATCTGGCTCGATCAGCGGGATTACGCGCGATACGCTGACTGACCGCAACA 5785  
Db |||||  
QY 3765 GCTGCGGACTGGCTGGCGAGCGG-----GGGACGGTTGAGGGTCCGGTGCC 3715  
Db |||||  
QY 5786 TCGTGAGGGCTCGCTCACTGCGAGCGGCTTTAGTGGCGGAGCGGAACCTATTATCA 5845  
Db |||||  
QY 3714 TTGGAGCGGCTTATTCCGTCATCATCGCCTGCTTGACGCGCGAGCCCTATTATGC 3655  
Db |||||  
QY 5846 CTCGGCGGATCCGGTTTACCGGGTTTACCGCTGCGCTCGGCTGGGCCATCGCTACGGC 5905  
Db |||||  
QY 3654 CTCGGCGGCGAGGGCTTCGGCATCTGCGCGCGCTGCGCTGCTGATCGCGCCGC 3595  
Db |||||  
QY 5906 TCGCGGGTTTATCGGAAATTGGCGTCAAAGTTGACGACCGCGTGTGCACGCTGGGA 5965  
Db |||||  
QY 3594 GCTGCGTATCTATCGCGCAATCGGACGCGCATCCGGCAGGGTGGCCCGAGGCTATCG 3535  
Db |||||  
QY 5966 TTCACGGCAGCGCACCAAGTAAAGGTGAACTGGGCGCTGCTGGTGAAGGGGCGAGTTT 6025  
Db |||||  
QY 3534 CCAGCGGATCAGCACGTCGRAGGCTGCCAAGATCGGGCTTCTGGCGCGGAGGCTTGA 3475  
Db |||||  
QY 6026 GCGGATCACTTCGCGTGTGTCTGTCCTGAACCGCGTCCGGGTGGTCTGTGGCAGCGTCC 6085  
Db |||||  
QY 3474 CGCGGCGCATCGCGCTGCGCGCGCGGAAATCAGCGGAGCGCTGTGGACCGGACG 3415  
Db |||||  
QY 6086 TCG 6088  
Db |||||  
QY 3414 GCG 3412

Search completed: January 26, 2005, 02:16:28  
Job time : 3368 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 24, 2005, 22:53:11 ; Search time 19359 Seconds  
(without alignments)  
13174.321 Million cell updates/sec

Title: US-10-808-979-18  
Perfect score: 6999  
Sequence: 1 ggcagggcaatggaaaatgt.....gacgttttagacatgcaata 6999

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

- 1: gb\_est1:\*
- 2: gb\_est2:\*
- 3: gb\_hic:\*
- 4: gb\_est3:\*
- 5: gb\_est4:\*
- 6: gb\_est5:\*
- 7: gb\_est6:\*
- 8: gb\_gss1:\*
- 9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	231.6	3.3	659	8 AQ989968	AQ989968 Rfc00652
C 2	131.4	1.9	621	8 AQ990546	AQ990546 Rfc01339
C 3	77.6	1.1	315	6 CB041878	CB041878 4006842 B
C 4	77.6	1.1	315	6 CB042040	CB042040 4007005 B
C 5	77.6	1.1	349	6 CB041738	CB041738 4006549 B
C 6	77.6	1.1	359	6 CB041616	CB041616 4006367 B
C 7	77.4	1.1	314	6 CB061284	CB061284 4011834 B
C 8	75.8	1.1	300	6 CB061218	CB061218 4011666 B
C 9	75.8	1.1	317	6 CB059804	CB059804 4009880 B
C 10	74.4	1.1	354	6 CB041601	CB041601 4006361 B
C 11	74.2	1.1	302	6 CB062064	CB062064 4012128 B
C 12	73.6	1.1	475	8 BH439236	BH439236 BOHAC18TF
C 13	73.6	1.1	754	8 BZ434062	BZ434062 BONF034TF
C 14	72	1.0	425	2 BE323612	BE323612 NF006A09P
C 15	72	1.0	709	8 BH502634	BH502634 BOHJK46TR
C 16	72	1.0	736	6 CA916997	CA916997 EST641144
C 17	71.8	1.0	775	8 BZ550668	BZ550668 Pacsl-60
C 18	71.6	1.0	809	7 CF517921	CF517921 CAP0005_I
C 19	71.6	1.0	833	7 CF513004	CF513004 CABud0005
C 20	71.2	1.0	632	7 CN144185	CN144185 WOUND1_20
C 21	71.2	1.0	725	7 CN144663	CN144663 WOUND1_23
C 22	70.8	1.0	1218	8 BZ575772	BZ575772 msh2 462.
C 23	70.4	1.0	682	5 BQ856404	BQ856404 QGB3G04.Y
C 24	70	1.0	453	1 AI442185	AI442185 sds8c11.Y

25	69.6	1.0	698	7	CN148357	CN148357 WOUND1_55
26	69.6	1.0	798	8	BZ563162	BZ563162 pacs2_164
C 27	69.2	1.0	761	9	CC648441	CC648441 OGUJW19TH
28	69.2	1.0	777	9	CC722834	CC722834 OGUJW11TH
C 29	69.2	1.0	807	9	CC648445	CC648445 OGUJW19TV
30	69.2	1.0	929	9	CG441377	CG441377 OGVHT32TV
31	69.2	1.0	1090	8	BZ553098	BZ553098 pacs1-60
32	68.6	1.0	666	4	BI717821	BI717821 1031021A0
33	68.6	1.0	675	4	BI723806	BI723806 1031067H0
34	68.6	1.0	701	2	BE238154	BE238154 894040E07
C 35	68.4	1.0	601	6	CB042842	CB042842 4007706 B
36	68.4	1.0	744	9	CC602991	CC602991 OGUJB14TH
37	68.4	1.0	785	9	CC661270	CC661270 OGMA033TH
C 38	68.2	1.0	771	8	BZ442823	BZ442823 BONH857TR
39	68.2	1.0	790	7	CF443389	CF443389 EST679734
40	67.6	1.0	824	6	CB066639	CB066639 EST645262
41	67.4	1.0	596	8	AQ963747	AQ963747 LRRGP06TF
42	67	1.0	668	4	BI726915	BI726915 1031088E0
43	66.2	0.9	947	7	CK295442	CK295442 EST758156
C 44	66.2	0.9	975	7	CK295932	CK295932 EST758646
C 45	66	0.9	809	7	CK285594	CK285594 EST748316

ALIGNMENTS

RESULT 1  
AQ989968/c  
LOCUS  
DEFINITION  
Rfc00652 Photorhabdus luminescens strain W14 M13 library  
Photobacterium luminescens genomic clone PLG00652, genomic survey  
sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT

AQ989968 659 bp DNA linear GSS 14-AUG-2000  
Rfc00652 Photorhabdus luminescens strain W14 M13 library  
Photobacterium luminescens genomic clone PLG00652, genomic survey  
sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT

Contact: ffrrench-Constant RH  
Department of Biology and Biochemistry  
University of Bath  
South Building, Bath BA2 7AY, UK  
Tel: (44) 1225 826621  
Fax: (44) 1225 826779  
Email: bssrfc@bath.ac.uk  
This is one of 2,122 random reads from the M13 library. For  
annotation of identified clones (BLASTX, BLASTN and mapping to E.  
coli K12 genome) please see ffrrench-Constant et al. 2000, Nucleic  
Acids Res.  
Seq primer: M13 Forward  
Class: shotgun.  
Location/Qualifiers  
1. .859  
/organism="Photobacterium luminescens"  
/mol\_type="genomic DNA"  
/strain="W14"  
/db\_xref="taxon:29488"  
/clone="PLG00652"  
/dev\_stage="primary phase variant"  
/clone\_lib="Photobacterium luminescens strain W14 M13  
library"  
/note="Genomic DNA from strain W14 was size selected (1-2  
kb) and then cloned into M13 Janus."

FEATURES  
source

ORIGIN



DEFINITION 4006842 BARC-EMBRAPA 315BOV Bos indicus cDNA clone 315BOV\_1E16  
 Unknown, mRNA sequence.  
 ACCESSION CB041878  
 VERSION CB041878.1 GI:27761123  
 KEYWORDS EST.  
 SOURCE Bos indicus (zebu)  
 ORGANISM Bos indicus (zebu)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.  
 REFERENCE 1 (bases 1 to 315)  
 AUTHORS da Mota, A.F., Sonstegard, T.S., Van Tassell, C.P., Matukumalli, L.K., Wood, D.L., Capuco, A.V., Brito, M.A.P., Martinez, M.L., Connor, E.E., Machado, M.A. and Coutinho, L.L.  
 TITLE Construction and Characterization of cDNA Libraries Generated from Mammary Gland Tissues of Holstein (Bos taurus) and Gir (Bos indicus) Cattle  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Adilson F. da Mota  
 Gene Evaluation and Mapping Laboratory  
 USDA, ARS, Animal and Natural Resources Institute  
 Bldg. 200 Rm3 BARC-East, Beltsville, MD 20705, USA  
 Tel: 3015048456  
 Fax: 3015048414  
 Email: amota@cnpgl.embrapa.br  
 Single pass sequencing. Bases called and trimmed with phred 0.000925 using options -trim alt - -trim fasta. Vector identified by cross match using options -minmatch 12 -minscore 18  
 PCR Primers  
 FORWARD: GTTTTCCAGTCACGAGTTG  
 BACKWARD: TGAGCGGATACAAATTCACACG  
 Plate: 1 row: E column: 16  
 Seq primer: GTTTTCCAGTCACGAGTTG  
 High quality sequence stop: 315.  
 FEATURES  
 source  
 1. .315  
 /organism="Bos indicus"  
 /mol\_type="mRNA"  
 /strain="Brazilian Dairy Gir"  
 /db\_xref="taxon:9915"  
 /clone="315BOV\_1E16"  
 /sex="female"  
 /tissue\_type="mammary"  
 /cell\_type="epithelium"  
 /dev\_stage="involved"  
 /lab\_host="K-12"  
 /clone\_lib="BARC-EMBRAPA 315BOV"  
 /note="Organ: mammary; Vector: pUC118; Site 1: HincII; Site 2: HincII; This mammary-derived cDNA library was created as part of a collaborative project between the ARS Gene Evaluation and Mapping Laboratory and the EMBRAPA Dairy Cattle Research Center under the sponsorship of USDA, ARS/EMBRAPA-LABEX program in animal genomics. RNA extracted on 6/8/02, RT with Superscript II at 37 deg C annealing temperature, PCR with 16mer AM5."  
 ORIGIN  
 Query Match 1.1%; Score 77.6; DB 6; Length 315;  
 Best Local Similarity 56.2%; Pred. No. 2.5e-09;  
 Matches 146; Conservative 0; Mismatches 114; Indels 0; Gaps 0;  
 QY 6313 CGTCATTCGGGACCAATCCATCGTGGAGGCGCCATGCTTGGCACACCGGACATAATCG 6372  
 DB 306 CGCACTTCCATCGCGATCGGAGCATCTCAAGGCCAGACGTGCAACCCGAGAGCCACAAA 247  
 QY 6373 CACAATTCGGCAGTACCGCAACACCCGCGCATAAAGATCGTTGAGTCTCAACTTCACGCT 6432  
 DB 246 GATGATGCGCGGGCGCGGAAACCCGCGGAACAGCGTCTTCTTCAGCAGCGTCGTC 187  
 QY 6433 GTCGGGTTTCATGTTGGCACAGATGCCAGGCGCCATCCCGGATCGATGATGATTTATG 6492  
 DB 186 GTGGGGCTCATGATGGTGGGTGCCAGGTCCACAGCAGGCGCGTGCATGACATAACGATG 127

QY 6493 CGACAGCGCGCTACGATTTCATCATCACACCGGTTGCTCCACACAGATAGCAGCTCCA 6552  
 DB 126 CATCGCCCAAGCGAGAGACCTCCATCCCGAGGAGTGTCGCGAGGACAGCAGGATCATCGT 67  
 QY 6553 TAACCAAGACGATTTGTTGCTC 6572  
 DB 66 CGGCCAGGACATGACCCATC 47  
 RESULT 4  
 CB042040/c  
 LOCUS CB042040  
 DEFINITION 4007005 BARC-EMBRAPA 315BOV Bos indicus cDNA clone 315BOV\_1P23  
 Unknown, mRNA sequence.  
 ACCESSION CB042040  
 VERSION CB042040.1 GI:27761285  
 KEYWORDS EST.  
 SOURCE Bos indicus (zebu)  
 ORGANISM Bos indicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.  
 REFERENCE 1 (bases 1 to 315)  
 AUTHORS da Mota, A.F., Sonstegard, T.S., Van Tassell, C.P., Matukumalli, L.K., Wood, D.L., Capuco, A.V., Brito, M.A.P., Martinez, M.L., Connor, E.E., Machado, M.A. and Coutinho, L.L.  
 TITLE Construction and Characterization of cDNA Libraries Generated from Mammary Gland Tissues of Holstein (Bos taurus) and Gir (Bos indicus) Cattle  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Adilson F. da Mota  
 Gene Evaluation and Mapping Laboratory  
 USDA, ARS, Animal and Natural Resources Institute  
 Bldg. 200 Rm3 BARC-East, Beltsville, MD 20705, USA  
 Tel: 3015048456  
 Fax: 3015048414  
 Email: amota@cnpgl.embrapa.br  
 Single pass sequencing. Bases called and trimmed with phred 0.000925 using options -trim alt - -trim fasta. Vector identified by cross match using options -minmatch 12 -minscore 18  
 PCR Primers  
 FORWARD: GTTTTCCAGTCACGAGTTG  
 BACKWARD: TGAGCGGATACAAATTCACACG  
 Plate: 1 row: P column: 23  
 Seq primer: GTTTTCCAGTCACGAGTTG  
 High quality sequence stop: 315.  
 FEATURES  
 Location/Qualifiers  
 1. .315  
 /organism="Bos indicus"  
 /mol\_type="mRNA"  
 /strain="Brazilian Dairy Gir"  
 /db\_xref="taxon:9915"  
 /clone="315BOV\_1P23"  
 /sex="female"  
 /tissue\_type="mammary"  
 /cell\_type="epithelium"  
 /dev\_stage="involved"  
 /lab\_host="K-12"  
 /clone\_lib="BARC-EMBRAPA 315BOV"  
 /note="Organ: mammary; Vector: pUC118; Site 1: HincII; Site 2: HincII; This mammary-derived cDNA library was created as part of a collaborative project between the ARS Gene Evaluation and Mapping Laboratory and the EMBRAPA Dairy Cattle Research Center under the sponsorship of USDA, ARS/EMBRAPA-LABEX program in animal genomics. RNA extracted on 6/8/02, RT with Superscript II at 37 deg C annealing temperature, PCR with 16mer AM5."  
 ORIGIN  
 Query Match 1.1%; Score 77.6; DB 6; Length 315;  
 Best Local Similarity 56.2%; Pred. No. 2.5e-09;  
 Matches 146; Conservative 0; Mismatches 114; Indels 0; Gaps 0;





Single pass sequencing. Bases called and trimmed with phred 0.000925 using options -trim\_alt -trim\_fasta. Vector identified by cross\_match using options -minmatch 12 -minscore 12

PCR Primers

FORWARD: GTTTCCTCCAGTCACGACGTTG

BACKWARD: TGAGCGGATACAAATTTACACAG

Plate: 1 row: H column: 20

Seq primer: GTTTCCTCCAGTCACGACGTTG

High quality sequence stop: 314.

Location/Qualifiers

1. 314

/organism="Bos indicus"

/mol\_type="mRNA"

/strain="Brazilian Dairy Gir"

/db\_xref="taxon:9915"

/clone="338BOV\_1H20"

/sex="female"

/tissue\_type="teat cistern and alveoli"

/cell\_type="multiple"

/dev\_stage="involved"

/lab\_host="K-12"

/clone\_lib="BARC-EMBRAPA 338BOV"

/note="Organ: mammary; Vector: pUC 118; Site 1: HincII;

Site 2: HincII; This mammary-derived cDNA library was

created as part of a collaborative project between the ARS

Gene Evaluation and Mapping Laboratory and the EMBRAPA

Dairy Cattle Research Center under the sponsorship of

USDA, ARS/EMBRAPA-LABEX program in animal genomics. RNA

extracted on 6/8/02, RT with Superscript II at 37 deg C

annealing temperature. Cow AM1."

# FEATURES

source

## ORIGIN

Query Match 1.1%; Score 77.4; DB 6; Length 314;  
Best Local Similarity 57.1%; Pred. No. 2.8e-09;  
Matches 141; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

6326 CCAATTCACATGCGAGCGCCACATGCTTGCACACCGACATTAATCAGCAATCGCCAGT 6385  
4 CCGATTCGCTGTCATCCAGGCGCAGAGCTGCAACCCGAGAGCCACAAAGATGATGGCGGG 63

6386 ACCGCAACACACCGCATTAAGATCGTTGAGCTCAAACTTACCGCTGTGCGGTTTCATGG 6445  
64 GGGCGGAAACACCGCGAAGAGTCTTCTTCACGACGCTGCTGCGGCTCATGA 123

6446 TCGCAGATGCGAGCGCCCATCCCAACCGTCGATGATGATTTATGCGACGCGCGCT 6505  
124 TGGTCGCTGCGAGCTCCACAGCGCGTGCATGACATACGATGCATCGCCATCG 183

6506 AGCATTTCCATCACCACCGAGTTCGCAACAGATAGCAGCTTCCATACCGAGCATTT 6565  
184 AAGACCTCCATCCCGAGGAGTGGCGAGGACAGCAGGATCATCTGCGCCAGGACATG 243

6566 GTTCGTC 6572

244 ACCCATC 250

## RESULT 8

CB061218/c

LOCUS

DEFINITION

Unknown, mRNA sequence.

ACCESSION

CB061218

VERSION

CB061218.1 GI:27799505

KEYWORDS

SOURCE

ORGANISM

Bos indicus (zebu)

Bos indicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Bovinae; Bos.

REFERENCE

1 (bases 1 to 300)

AUTHORS

da Mota,A.F., Sonstegard,T.S., Van Tassell,C.P., Matukumalli,L.K.,

Wood,D.L., Capuco,A.V., Brito,M.A.P., Martinez,M.L., Connor,E.E.,

/mol\_type="mRNA"  
/strain="Brazilian Dairy Gir"  
/db\_xref="taxon:9915"  
/clone="314BOV\_1F09"  
/sex="female"  
/tissue\_type="teat cistern"  
/cell\_type="epithelium"  
/dev\_stage="involved"  
/lab\_host="K-12"  
/clone\_lib="BARC-EMBRAPA 314BOV"  
/note="Organ: mammary; Vector: pUC118; Site 1: HincII;  
Site 2: HincII; This mammary-derived cDNA library was  
created as part of a collaborative project between the ARS  
Gene Evaluation and Mapping Laboratory and the EMBRAPA  
Dairy Cattle Research Center under the sponsorship of  
USDA, ARS/EMBRAPA-LABEX program in animal genomics. RNA  
extracted on 6/8/02, RT with Superscript II at 37 deg C  
annealing temperature. PCR with 16mer AMS only."

## ORIGIN

Query Match 1.1%; Score 77.6; DB 6; Length 359;  
Best Local Similarity 56.2%; Pred. No. 2.6e-09;  
Matches 146; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

6313 CGTCATTCGGCACCACATCCTGCGAGCGCCACATGCTTGCACACCGACATTAATCAG 6372  
20 CGCACTTCCATCCGATCGGCAGCATCCAAAGCCAGAGCTGCAACCCGAGAGCCACAAA 79

6373 CACAATCGCCAGTACCGCAACACACCGCATTAAGATGTTGAGCTCAAACTTACCGCT 6432  
80 GATGATGCGCGGGCGGGAACACCGCGAAGAGTCTTCTTCACGACGCTGCTC 139

6433 GTCCGGTTTCATGTCGACAGATGCCAGCGCCATCCCAACCGTCGATGATGATTTATG 6492  
140 GTCCGGTTCATGATGCTGCGGTGCCAGGTCACAGCGCGCTGCATGACATACGATG 199

6493 CGACAGCGCGCTACGATTTCCATCACCACCGTTCGCAACAGATAGCAGTTCCA 6552  
200 CATCGCCCAAGCAAGACCTTCATCCCGAGGAGTGGCGAGGACAGCAGGATCATCGT 259

6553 TAACCAAGACATGTTGCTC 6572

260 CGGCCAGGATACCCATC 279

## RESULT 7

CB061284

LOCUS

DEFINITION

Unknown, mRNA sequence.

ACCESSION

CB061284

VERSION

CB061284.1 GI:27799571

KEYWORDS

SOURCE

ORGANISM

Bos indicus (zebu)

Bos indicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Bovinae; Bos.

REFERENCE

1 (bases 1 to 314)

AUTHORS

da Mota,A.F., Sonstegard,T.S., Van Tassell,C.P., Matukumalli,L.K.,

Wood,D.L., Capuco,A.V., Brito,M.A.P., Martinez,M.L., Connor,E.E.,

Machado,M.A. and Coutinho,L.B.

Construction and Characterization of cDNA Libraries Generated from

Mammary Gland Tissues of Holstein (Bos taurus) and Gir (Bos

indicus) Cattle

Unpublished (2002)

Contact: Adilson F. da Mota

Gene Evaluation and Mapping Laboratory

USDA, ARS, Animal and Natural Resources Institute

Bldg. 200 Rm3 BARC-East, Beltsville, MD 20705, USA

Tel: 3015048456

Fax: 3015048414

Email: amota@cnpgl.embrapa.br



QY 6506 ACATTTCATCACCACCGGTTGCCACAGATAGACAGTTCATACACAGACATT 6565  
 Db 131 AAGACTTCATCCCAAGGAGGTGGCGAGGAACAGCAGGATCGTGTGGCCAGGACATG 72  
 QY 6566 GTTCGTC 6572  
 Db 71 ACCATC 65

RESULT 10  
 CB041601/c 354 bp mRNA linear EST 15-JAN-2003  
 LOCUS 4006361 BARC-EMBRAPA 314BOV Bos indicus cDNA clone 314BOV\_1E09  
 DEFINITION Unknown, mRNA sequence.  
 ACCESSION CB041601  
 VERSION CB041601.1 GI:27760846  
 KEYWORDS EST.  
 SOURCE Bos indicus (zebu)  
 ORGANISM Bos indicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 Bovinae; Bos.

REFERENCE 1 (bases 1 to 354)  
 AUTHORS da Mota,A.F., Sonstegard,T.S., Van Tassel,C.P., Matukumalli,L.K.,  
 Wood,D.L., Capuco,A.V., Brito,M.A.P., Martinez,M.L., Connor,E.E.,  
 Machado,M.A. and Coutinho,L.L.  
 TITLE Construction and Characterization of cDNA Libraries Generated from  
 Mammary Gland Tissues of Holstein (Bos taurus) and Gir (Bos  
 indicus) Cattle

JOURNAL Unpublished (2002)

COMMENT Contact: Adilson F. da Mota

Gene Evaluation and Mapping Laboratory

USDA, ARS, Animal and Natural Resources Institute

Bldg. 200 Rm3 BARC-East, Beltsville, MD 20705, USA

Tel: 3015048456

Fax: 3015048414

Email: amota@npgl.embrapa.br

Single pass sequencing. Bases called and trimmed with phred  
 0.000925 using options -trim alt -trim fasta. Vector identified  
 by cross match using options -minmatch 12 -minscore 12

PCR Primers

FORWARD: GTTTCCTCCAGTCACGACGTTG

BACKWARD: TGACGGGATACAAATTTCACAG

Plate: 1 row: E column: 09

Seq primer: GTTTCCTCCAGTCACGACGTTG

High quality sequence stop: 354.

FEATURES

source

1..354

/organism="Bos indicus"

/mol\_type="mRNA"

/strain="Brazilian Dairy Gir"

/db\_xref="taxon:9915"

/clone="314BOV\_1E09"

/sex="female"

/tissue\_type="teat cistern"

/cell\_type="epithelium"

/dev\_stage="involved"

/lab\_host="K-12"

/clone\_lib="BARC-EMBRAPA 314BOV"

/note="Organ: mammary; Vector: pUC118; Site 1: HincII;  
 Site 2: HincII; This mammary-derived cDNA library was  
 created as part of a collaborative project between the ARS  
 Gene Evaluation and Mapping Laboratory and the EMBRAPA  
 Dairy Cattle Research Center under the sponsorship of  
 USDA, ARS/EMBRAPA-LABEX program in animal genomics. RNA  
 extracted on 6/8/02, RT with Superscript II at 37 deg C  
 annealing temperature, PCR with 16mer AWS only."

ORIGIN

Query Match 1.1%; Score 74.4; DB 6; Length 354;

Best Local Similarity 55.4%; Pred. No. 1.9e-08;

Matches 144; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 6313 CGTCTTCGGCACCACCAATCCATCTGCAGCGGCACATGCTTGCACACCGACATAATCAG 6372  
 Db 345 CGCACTTCCCATCGCCGATCGGAGCATCCAAGCCACAGCGTGCACCCGAGAGCCACAA 286  
 QY 6373 CACAATCGCAGTACCGCAAAACACACCGCATATAAGATCGTTGAGCTCAAACCTTACCGCT 6432  
 Db 285 GATGATGGCGGGGGCGGGAACACACCGGACAGGTGCTCTTCTCCAGCAGCTCGTC 226  
 QY 6433 GTCCGCTTCATGTCGACAGATGCCAGCCCATCCCCACCGTGCATGATGATTTATG 6492  
 Db 225 GTGCGGCTCATGATGCTGCGGTGCCAGGTCCACAGCAGCGCTGCATGACATAACGATG 166  
 QY 6493 CGACAGCGCGCTACGATTTCCATCACACACCGTTCACACAGATAGCAGCTTCCA 6552  
 Db 165 CATGCCCCAGCGAAGACCTCGTCCCCAGGAGGTGGCGGAGAACAGCAGGATCATCGT 106  
 QY 6553 TAAACGAGACATTTGTTGCTC 6572  
 Db 105 CGGCCAGGACATGACCCATC 86

RESULT 11  
 CB062064/c

LOCUS 4012128 BARC-EMBRAPA 339BOV Bos indicus cDNA clone 339BOV\_1114

DEFINITION Unknown, mRNA sequence.

ACCESSION CB062064

VERSION CB062064.1 GI:27800351

KEYWORDS EST.

SOURCE Bos indicus (zebu)

ORGANISM Bos indicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Bovinae; Bos.

REFERENCE 1 (bases 1 to 302)

AUTHORS da Mota,A.F., Sonstegard,T.S., Van Tassel,C.P., Matukumalli,L.K.,

Wood,D.L., Capuco,A.V., Brito,M.A.P., Martinez,M.L., Connor,E.E.,

Machado,M.A. and Coutinho,L.L.

TITLE Construction and Characterization of cDNA Libraries Generated from

Mammary Gland Tissues of Holstein (Bos taurus) and Gir (Bos

indicus) Cattle

JOURNAL Unpublished (2002)

COMMENT Contact: Adilson F. da Mota

Gene Evaluation and Mapping Laboratory

USDA, ARS, Animal and Natural Resources Institute

Bldg. 200 Rm3 BARC-East, Beltsville, MD 20705, USA

Tel: 3015048456

Fax: 3015048414

Email: amota@npgl.embrapa.br

Single pass sequencing. Bases called and trimmed with phred

0.000925 using options -trim alt -trim fasta. Vector identified

by cross match using options -minmatch 12 -minscore 12

PCR Primers

FORWARD: GTTTCCTCCAGTCACGACGTTG

BACKWARD: TGACGGGATACAAATTTCACAG

Plate: 1 row: 1 column: 14

Seq primer: GTTTCCTCCAGTCACGACGTTG

High quality sequence stop: 302.

FEATURES

source

1..302

/organism="Bos indicus"

/mol\_type="mRNA"

/strain="Brazilian Dairy Gir"

/db\_xref="taxon:9915"

/clone="339BOV\_1114"

/sex="female"

/tissue\_type="teat cistern and alveoli"

/cell\_type="multiple"

/dev\_stage="involved"

/lab\_host="K-12"

/clone\_lib="BARC-EMBRAPA 339BOV"

/note="Organ: mammary; Vector: pUC 19; Site 1: SmaI;

Site 2: SmaI; This mammary-derived cDNA library was

created as part of a collaborative project between the ARS Gene Evaluation and Mapping Laboratory and the EMBRAPA Dairy Cattle Research Center under the sponsorship of USDA, ARS/EMBRAPA-LABEX program in animal genomics. RNA extracted on 6/8/02, RT with Superscript II at 37 deg C annealing temperature. Cow AWL."

## ORIGIN

Query Match 1.1%; Score 74.2; DB 6; Length 302;  
Best Local Similarity 56.3%; Pred. No. 2.1e-08;  
Matches 139; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 6326 CCAATCCATGCGAGCGGCACATGCTTGACACCGACATAAATCAGCAATCGCCAT 6385  
DB 296 CCGATCGGAGCATCAAGCGCCAGATGTGCAACCCGAGAGCCACAAGATGATGCGCGG 237

QY 6386 ACCGCAACACCCACCGCATTAAGATCGTTGAGCTCAAACTTACCGCTGTCGGTTTCATGG 6445  
DB 236 GCGGCGAAACACCCGCAACAGGTCTTCTTCAGCACCGTCGTCGCGGCTCATGA 177

QY 6446 TGGCAGCATGCCAGCGCCCATCCCAACCGTGATGATGTATTATGCGACAGCGCGCT 6505  
DB 176 TGGCTCGGTGCGAGTTCACAGCAGCGCGTGATGACATACGATGATCGCCCAAGCG 117

QY 6506 AGATTTCATCACCAACCGTTCGCAACAGATGAAGCACCGTTCCATTAACGAGCAT 6565  
DB 116 AAGACCTCCATCCCGAGGAGTGGCGAGACAGCAGGATCATCGTCGCCAGGACATG 57

QY 6566 GTTCGTC 6572  
DB 56 ACCATC 50

## RESULT 12

BH439236/c  
LOCUS BH439236 475 bp DNA linear GSS 12-DEC-2001  
DEFINITION BOHAC18TF BOHA Brassica oleracea genomic clone BOHAC18, genomic survey sequence.

ACCESSION BH439236  
VERSION BH439236.1 GI:17624950  
KEYWORDS GSS.  
SOURCE Brassica oleracea  
ORGANISM Brassica oleracea  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE 1 (bases 1 to 475)  
TOWN,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.  
Whole genome shotgun sequencing of Brassica oleracea  
Unpublished (2001)  
Other GSSs: BOHAC18TR  
Contact: Chris Town

TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TF  
Class: sheared ends.

## FEATURES

source  
1..475  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="TO1000DH3"  
/db\_xref="taxon:3712"  
/clone\_lib="BOHA"  
/note="Vector: PHOS1; Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into PHOS1 using BstXI linkers"

## ORIGIN

Query Match 1.1%; Score 73.6; DB 8; Length 475;  
Best Local Similarity 53.9%; Pred. No. 3.9e-08;  
Matches 151; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY 559 GCCTGTGAGTGAATGTCACCGGCGATCGCTGATTTGGATGACATTCCTCGATG 618  
DB 194 GCTTGGCGGTGAAATGATTCACACGATGTCGCTGATCCACGACGATCTCCCGTATG 253

QY 619 GATACGCGCAGATCGTGTGTCGCCCTACCGTGCAGTAATTTGGTGAACACG 678

Best Local Similarity 53.9%; Pred. No. 3.5e-08;  
Matches 151; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY 559 GCCTGTGAGTGAATGTCACCGGCGATCGCTGATTTGGATGACATTCCTCGATG 618  
DB 378 GCTTGGCGGTGAAATGATTCACACGATGTCGCTGATCCACGACGATCTCCCGTATG 319

QY 619 GATAACGCGCAGATCGTGTGTCGCCCTACCGTCATCGCGAATTTGGTGAACACG 678  
DB 318 GACAAGCAGACCTCGCGCGGAAACCCACCAACCAAGTGTTCGGCGAAGACGTC 259

QY 679 GCGATTTCGCGCCCATCGCTGTAGCCGCGGATTTGAAGTGAATTCCTCGATTCGACCC 738  
DB 258 GCGGTTTTAGCCGAGACGCGCTTTTATCGTTTCGCTTGAGCACTTGGCGTCGACGTCG 199

QY 739 GCTTTCCTGCCATACATAAATCTGAAGCATTCGTAACCTCCGCTGCCGTCGCGCTG 798  
DB 198 GGGGTGGCTTCGCGCGAGGGTGTAGACGATCGGGGAGCTGGCGAGAGCCGTTGGATCA 139

QY 799 CAGGCGCTTAGTCGAAGGGCAATTCAGGATCTGCACGACG 838  
DB 138 GAAGGGCTTGTGGCGGTTCAGATTGTGGATATCAGCAGCG 99

## RESULT 13

BZ434062  
LOCUS BZ434062 754 bp DNA linear GSS 13-DEC-2002  
DEFINITION BONFQ34TF BO 1.6.2\_KB tot Brassica oleracea genomic clone BONFQ34, genomic survey sequence.

ACCESSION BZ434062  
VERSION BZ434062.1 GI:26683951  
KEYWORDS GSS.  
SOURCE Brassica oleracea  
ORGANISM Brassica oleracea  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE 1 (bases 1 to 754)  
TOWN,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.  
Whole genome shotgun sequencing of Brassica oleracea  
Unpublished (2001)  
Other GSSs: BONFQ34TR  
Contact: Chris Town

TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TF  
Class: sheared ends.

## FEATURES

source  
1..754  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="TO1000DH3"  
/db\_xref="taxon:3712"  
/clone="BONFQ34"  
/clone\_lib="BO 1.6.2\_KB tot"  
/note="Vector: PHOS1; Site 1: BstXI; 1.6-2 kb sheared total DNA inserted into PHOS1 using BstXI linkers"

## ORIGIN

Query Match 1.1%; Score 73.6; DB 8; Length 754;  
Best Local Similarity 53.9%; Pred. No. 3.9e-08;  
Matches 151; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY 559 GCCTGTGAGTGAATGTCACCGGCGATCGCTGATTTGGATGACATTCCTCGATG 618  
DB 194 GCTTGGCGGTGAAATGATTCACACGATGTCGCTGATCCACGACGATCTCCCGTATG 253

QY 619 GATACGCGCAGATCGTGTGTCGCCCTACCGTGCAGTAATTTGGTGAACACG 678

```

Db      254  GACAACGACGACTCCGCGCGGAAACCCACCAACCAACCAAGTGTTCGGCGAGACGTC 313
QY      679  GCGAATTCGCGCCCATCGCGTCTTAGCGCGCATTTGAAGTGAATGCCATTGCACCC 738
Db      314  GCGGTTTAAAGCGGAGACGCGCTTTTATCGCTTGGCTTGGAGCATTGGCGTGCAGCTCG 373
QY      739  GGTTCGCTGCCATACATAAATCTGAAGGATGCTGAACCTCTCCGCTCGCGTGCAGCTG 798
Db      374  GGGGTGGCTCCGCGGAGGCTGGTGTAGAGCATCGGGAGCTGGCGAGAGCCGCTGGATCA 433
QY      799  CAGGGCTTAGTGAAGGCAATTCAGGATCTGCAGCAG 838
Db      434  GAAGGCTTGTGCGGGTCAGATTGGATATCAGCAGCG 473

RESULT 14
LOCUS   BE323612
DEFINITION BE323612 425 bp mRNA linear EST 21-DRC-2000
VERSION   NF006A09PL1F1065 Phosphate starved leaf Medicago truncatula cDNA
KEYWORDS  clone NF006A09PL 5', mRNA sequence.
SOURCE   BE323612
ORGANISM BE323612.2 GI:11967172
          Medicago truncatula (barrel medic)
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
          rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
          Medicago.
REFERENCE 1 (bases 1 to 425)
AUTHORS   Liu,J., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
          Flores,H.R., Inman,J.T., Weller,J.W., May,G.D. and Harrison,M.J.
TITLE     Expressed Sequence Tags from the Samuel Roberts Noble Foundation
          Medicago truncatula phosphate-starved leaf library
JOURNAL   Unpublished (2000)
COMMENT   On Jul 14, 2000 this sequence version replaced gi:9197389.
          Contact: Harrison MJ
          Plant Biology Division
          The Samuel Roberts Noble Foundation
          2510 Sam Noble Parkway, Ardmore, OK 73402, USA
          Tel: 580 221 7325
          Fax: 580 221 7380
          Email: mjharrison@noble.org
          Medicago Genome Initiative accession: MGI:S:20177
          Insert Length: 818 Std Error: 0.00
          Plate: 006 row: A column: 09
          Seq primer: TCACACAGGAACACAGCTATGAC.
FEATURES
          Location/Qualifiers
            1..425
              /organism="Medicago truncatula"
              /mol_type="mRNA"
              /db_xref="taxon:3880"
              /clone="NF006A09PL"
              /tissue_type="leaf"
              /dev_stage="trifoliolate"
              /clone_lib="Phosphate starved leaf"
              /note="Vector: Lambda Zap; At the trifoliolate stage, M.
              truncatula plants were transplanted to phosphate-free sand
              and grown for a further 30 days. During this 30 day
              period, the plants were fertilized twice weekly with 1/2
              Hoaglands solution containing only 200m potassium
              phosphate. RNA was prepared from above ground tissues."
ORIGIN
          Query Match 1.0%; Score 72; DB 2; Length 425;
          Best Local Similarity 62.0%; Pred. No. 9.2e-08;
          Matches 114; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY      558  CGCTGTGCAAGTGAATGGTGCACGCGCATCGCTGATTCGTGATGACATTCCTCCGAT 617
Db      233  CGCGCTCGCGGTGAATGATCCACAGATGCTCTCTCCACGATGACCTTCCTTGAT 292
QY      618  GGATAACGCGCAGATCGCTGCTGGTCCGCTACCGTGCATCGCAATTTGGTGAACGCT 677

```

```

Db      293  GGATAACGATGATCTCCGCGAGGTAAACCTACAAACCAAGTCTTCGGAGAACGCT 352
QY      678  GCGGATTCCTCGCGCCATCGCGTCTTAGCGCGCATTTGAAGTGAATGCCATTGCACCC 737
Db      353  TCGTGTCTTCGCGCGGAGATGCTCTTCTCGCTTTGCTTCGAACATATTGCCGTCTCCAC 412
QY      738  CGGT 741
Db      413  CGTT 416

RESULT 15
LOCUS   BH502634
DEFINITION BH502634 BOHJ Brassica oleracea genomic clone BOHJK46, genomic
          survey sequence.
ACCESSION BH502634
VERSION   BH502634.1 GI:17710731
KEYWORDS  GSS.
SOURCE   Brassica oleracea
ORGANISM Brassica oleracea
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
          rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 709)
AUTHORS   Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
TITLE     Whole genome shotgun sequencing of Brassica oleracea
JOURNAL   Unpublished (2001)
COMMENT   Other GSSs: BOHJK46TF
          Contact: Chris Town
          TIGR
          9712 Medical Center Drive, Rockville, MD 20850, USA.
          Tel: 301-838-3523
          Fax: 301-838-0208
          Email: cdtown@tigr.org
          DNA is from a doubled haploid provided by Tom Osborn.
          Seq primer: TR
          Class: sheared ends.
FEATURES
          Location/Qualifiers
            1..709
              /organism="Brassica oleracea"
              /mol_type="genomic DNA"
              /strain="T01000DH3"
              /db_xref="taxon:3712"
              /clone="BOHJK46"
              /clone_lib="BOHJ"
              /note="Vector: pHOS1; Site 1: BstXI; 2-3 kb sheared
              genomic DNA inserted into pHOS1 using BstXI linkers"
ORIGIN
          Query Match 1.0%; Score 72; DB 8; Length 709;
          Best Local Similarity 50.2%; Pred. No. 1.1e-07;
          Matches 239; Conservative 0; Mismatches 225; Indels 12; Gaps 2;

QY      559  GCGTGTGCAAGTGAATGGTGCACGCGCATCGCTGATTCGTGATGACATTCCTCGATG 618
Db      87   GCTTCGCGGTTCGATGATGATTCACACGATGTCATCATCAAGACGATCTTCTTCGATG 146
QY      619  GATAACGCGCAGATCGCTGCTGCTCGCTTACCGTTCATCGGATTCGGTGAACCGTG 678
Db      147  GACAACGACGACCTCCGCGAGGTAAAGCCACGACGACACAAAGTCTTCGGAGAACGCTC 206
QY      679  GCGATTCTCGCGCCATCGCGTCTTAGCCCGCATTTGAAGTGAATTCGCAATTCGACCC 738
Db      207  GCGATTCTTCGCGCGCGCGCTCTTAGCTCTGGCTTCGACGATTTGACGAGTGCAC 266
QY      739  GGTTCGCTGCCATACATAAATCTGAAGCGATTCGTGAACCTCTCCGCTCGCGCTCG 798
Db      267  GTGTCGCTTGAGAGA---ATGTTAGAGCGGTAAAGAACTCGCGAAGTCTATAGGACG 323
QY      799  CAGGGCTTAGTCAGGGCAATTCAGGATCTGACGAC-----GGCAGCGACAGC 849

```

Db 324 AAAGGCTCGTGGCGGACAAGCGATGATTTGAGCAGCGAAGGTTTGGATCAAAACGAC 383  
QY 850 CGCAGCCCGAAGCGATCGCCATGACCAACGAACCTGAACACGAGGTGCTGTTTCGCGCC 909  
Db 384 GTCGGTTTGAAGGAGCTTGAGTTTATTCACGTTTATAAACCGGTTTCGCTGCTTGAGGCT 443  
QY 910 ACGCTGCAAAATGGCGGCGATTGCGGCTGACGCTTACCGCAGGTGCGGCAAGACTTAGC 969  
Db 444 TCGGCGTTATCGGAGCGGTTATTGGAGGTGTTTCGGAGGAAGAGGTTGAGAAGGTGAGG 503  
QY 970 TTCTTCGCCCCAGGATTTGGGCCAGGCGTTTCACTGCTCGACGACCTCGCCGACGG 1025  
Db 504 AGGTTCCGAGGTGATTGGGTTGTTTTCAGGTGTTGATGATATTTTGATGG 559

Search completed: January 25, 2005, 10:29:04  
Job time : 19364 secs